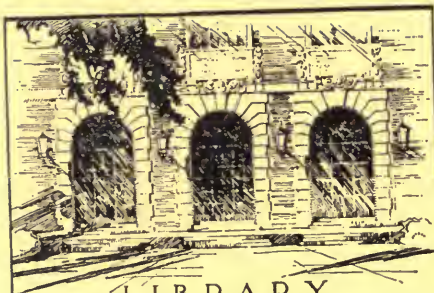


**PROCEEDINGS OF THE 1966 CLINIC
ON LIBRARY APPLICATIONS
OF DATA PROCESSING**

**Graduate School of Library Science
University of Illinois**



LIBRARY
OF THE
UNIVERSITY
OF ILLINOIS

025

C64p

1966

cop.3

The person charging this material is responsible for its return on or before the **Latest Date** stamped below.

Theft, mutilation, and underlining of books are reasons for disciplinary action and may result in dismissal from the University.

University of Illinois Library

JAN - 5 1969

L161—O-1096

UNIVERSITY OF ILLINOIS GRADUATE SCHOOL
OF LIBRARY SCIENCE

PROCEEDINGS OF THE 1966 CLINIC ON LIBRARY
APPLICATIONS OF DATA PROCESSING

Held at the Illini Union on the
Urbana Campus of the University
of Illinois, April 24-27, 1966

Edited by

HERBERT GOLDHOR

Distributed by
The Illini Union Bookstore
Champaign, Illinois

Copyright © 1966 by
The Board of Trustees of the University of Illinois

OTHER VOLUMES IN THIS SERIES

Proceedings of the 1963 Clinic on Library Applications of Data Processing. Pp. 176.

Proceedings of the 1964 Clinic on Library Applications of Data Processing. Pp. 117.

Proceedings of the 1965 Clinic on Library Applications of Data Processing. Pp. 201.

Available from the Illini Union Bookstore, 715 S. Wright St.,
Champaign, Ill. 61820. \$2.00 in paper covers, \$3.00 in cloth
covers.

TABLE OF CONTENTS

	Page
FOREWORD	v
CURRENT TRENDS IN LIBRARY AUTOMATION	
Joseph Becker	1
THE TECHNICAL INFORMATION PROJECT OF THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY	7
THE JOHNS HOPKINS UNIVERSITY LIBRARY	
Benjamin Courtright	18
THE USES OF PLATO: A COMPUTER CONTROLLED TEACHING SYSTEM	
Donald L. Bitzer, Elizabeth R. Lyman, and John A. Easley, Jr.	34
APPROACHES TO LIBRARY FILING BY COMPUTER	
Jean M. Perreault	47
INFORMATION RETRIEVAL PROJECTS IN THE BIOMEDICAL LIBRARY, UNIVERSITY OF CALIFORNIA, LOS ANGELES	
Louise Darling	91
OPERATIONAL COMPUTER-BASED SYSTEMS AT THE CHEMICAL ABSTRACTS SERVICE	
Seldon W. Terrant, Jr., and J. L. Wood	124
DATA PROCESSING IN THE TEXAS A & M LIBRARY	
Bruce W. Stewart	167
COMPUTER APPLICATIONS TO BOOK CATALOGS AND LIBRARY SYSTEMS	
Donald H. Stromberg	195
LIBRARY APPLICATIONS OF DATA PROCESSING: AN ENUMERATIVE BIBLIOGRAPHY, 1964-65	
James Krikelas	211

025
C64p
1966
cop. 3

FOREWORD

We present in this volume the papers given at the fourth annual Clinic on Library Applications of Data Processing, held at Urbana, Illinois, from April 24-27, 1966, under the sponsorship of the University of Illinois Graduate School of Library Science.

"The Use of a Remote Console in Programmed Teaching by Computer" is the title of the speech and the demonstration presented by Professor Donald L. Bitzer at the Clinic. As this speech was not recorded, we are reprinting, with the permission of the publisher, an article entitled "The Uses of PLATO, a Computer Controlled Teaching System," which appeared in the January 1966 issue of Audio Visual Instruction. This article by Professors Bitzer, Lyman, and Easley contains in essence the material presented by Donald Bitzer at the Clinic.

Not included in this volume are the remarks by Neil McMullan, Robert Trocchi, and Philip St. Aubin of the Minneapolis-Honeywell Company; they described the features and library uses of Honeywell computers, especially the 200 series. Also omitted from this publication are the discussions which followed each paper and the content of the problem session.

It seems clear that data processing has already affected many libraries and is likely to affect many more in the near future. The university and special libraries have been among the first to become involved, and usually at the computer level; school and public libraries are beginning to be involved, and usually at the level of punched cards. But involvement is only a first step along the road to successful operation of effective systems of data processing; to travel this road, librarians need to learn many new things.

In this process of successful application to libraries of data processing, the University of Illinois Graduate School of Library Science seeks to make a contribution by way of this annual conference or Clinic. The focus of the meeting is on the reporting of actual experience in case history form. We seek promising and fruitful experience (e.g., PLATO's use of a console-computer hook-up) and not necessarily successful experience alone; for surely we can learn much from mistakes and failures too. Libraries which have had experience with computers, which should be shared with others in the profession, are invited to communicate with the School for a possible place in future Clinics. Besides the emphasis on case reports, the

Clinic pattern includes each year a presentation by a manufacturer of data processing equipment, a paper on a functional technique (in this volume, it is Jean Perreault's paper "Approaches to Library Filing by Computer") and a general paper. This pattern seems to have been appropriate for the registrants at the Clinics so far to judge from their remarks and from the capacity audience we have usually had.

The planning committee for the 1966 Clinic consisted of Mrs. Frances B. Jenkins and Mr. Dewey E. Carroll of the Library School faculty. I wish to thank them and our other colleagues for their active assistance and participation in the Clinic. The speakers, of course, were the major contributors to whatever success the Clinic had, and all of us here appreciate their effort. The registrants too helped in the educative process, and their presence and cooperation were our reward.

Herbert Goldhor

Urbana, Illinois
November, 1966

CURRENT TRENDS IN LIBRARY AUTOMATION

Joseph Becker

The punched card is a remarkable invention which has evolved into a multi-billion dollar industry and has made its impact felt on almost every aspect of our society. The genesis of the punched card can be traced to the man credited with inventing it, Herman Hollerith. In the late 1800's, Hollerith cut a card to the exact dimensions of the American dollar bill and devised a method for representing a number or a letter in the identical place on each card every time. Although the dollar bill has shrunk, the Hollerith punched card, after a hundred years, has not changed by a millimeter! Because paper is a non-conducting material, it is possible to perform counting operations by passing electric current through the holes in the card. It was this simple idea which helped the United States analyze statistics collected by the 1890 census and which later led to many other applications, including some of interest to libraries.

Herman Hollerith's biography in the Dictionary of American Biography¹ relates that the punched card idea was suggested to him by a librarian. Hollerith thus reports the incident in one of his letters: "One evening at Dr. B's tea table he said to me, 'There ought to be a machine for doing the purely mechanical work of tabulating population and similar statistics.'" The "Dr. B" whom Hollerith refers to was Dr. John Shaw Billings, who was then Librarian of the Army Surgeon General's Library and who was destined to become the first Director of the New York Public Library. To this chance remark, Hollerith attributes his inspiration for the development of the punched card. Since it was a librarian who started it all, is it any wonder that Dr. Billings' professional descendants should wish to emulate his foresight by considering possible uses of data processing in libraries?

Over the years, punched cards were gradually applied to diverse areas of business, and by 1930 even universities were employing machines to perform functions associated with accounting and enrollment.

Joseph Becker is a Lecturer in the School of Library Science of Catholic University of America, Washington, D. C.

It was just about at this time that Ralph Parker, then a junior Librarian at the University of Texas, conceived of the idea of using punched card equipment for circulation work. The Director of the University of Texas Library was Don Coney, and Parker recalls, with good humor, how after many months of persuasion Coney finally gave him a \$300 grant for experimentation—but only after cautioning him to spend the money wisely! Another milestone in library punched card history was passed in the following decade when Marjorie Quigley, Librarian of the Montclair Public Library in New Jersey, acquired special-purpose equipment for controlling book transactions. This system of circulation control was the first to adopt the method of joining a machine-readable book card and a machine-readable borrower's card in a single master record at borrowing time. This twenty-five-year-old pilot punched card installation was the forerunner of the IBM 357 Data Collection System used by some libraries today for computerized circulation work.

The last twenty years have witnessed increasing interest among librarians concerning the possibility of using punched card machines and, more recently, computers, to carry out many other library functions. The reasons are clear. First, the rate of publishing has climbed steadily, dramatically increasing the number of printed pieces to be acquired, processed, housed, and circulated by libraries. Secondly, a rapidly expanding and more literate population has generated demands for reader services that have far exceeded a library's ability to respond effectively with traditional methods and techniques. Prospects for the future are even more staggering. Hence the professional librarian has been prompted to look for help to the new technology available in modern data processing equipment and systems.

Before the digital computer can be put to work constructively in libraries, its power and limitations must first be understood by the professional librarian. Data Processing Clinics, such as this one at the University of Illinois, and data processing courses, which are beginning to appear in library schools throughout the country, provide excellent opportunities for learning.

Computers, in the present day sense, first became known during the late 1940's. Ungainly, 30-ton, vacuum-tube heavyweights, they required their own power generators and air conditioning systems. Although their reliability left much to be desired, these first-generation computers were able to perform mathematical operations in a fraction of a second—a speed several orders of magnitude greater than was possible with the manual processing previously available. A second generation of transistorized computers became available in the late 1950's. Most of the machines we see around us today are the results of this metamorphosis. They are faster, cooler, and more accurate, and they possess versatility and personality far superior to

those of their predecessors. Furthermore, these machines no longer merely compute; they are capable of manipulating with equal skill letters of the alphabet, words, and sentences.

Many and varied computer programs have been written for second-generation machines to solve three classes of library problems. The first has to do with the use of computers for supporting the clerical functions found in technical processing and circulation work. Several libraries have had programs written which cause computers automatically to perform certain routine work, such as interfiling entries in a catalog, ordering books from publishers, writing request letters to the Library of Congress for cards, preparing serial records lists, monitoring circulation operations, printing book catalogs, and analyzing service to readers, among many other tasks. These applications, as in business, are designed to reduce the clerical burden, while at the same time increasing an organization's ability to perform more work.

A second category of computer programs that concern the librarian is the field of information storage and retrieval. Here the objective is to develop new intellectual methods for automatically extracting meaning from a text. If the text in question happens to be catalog data, the problem is simplified because the elements of bibliographic information are fixed and easily identifiable. However, if the purpose is to correlate facts or infer subject relationships from the complete content of articles and books, then the problem becomes much more complex. The latter area will require additional research before we can see clearly how such a capability will affect the duties of the reference librarian.

Until recently librarians paid scant attention to the application of mathematics and computers to the decision-making process in library management. This third area, referred to as operations research, is one which employs the principles of scientific management, but whereas scientific management is concerned with efficiency, operations research is concerned with effectiveness. At the turn of the century, the founder of scientific management, Dr. F. W. Taylor, advocated three things: (1) an inquiring frame of mind, which refuses to accept past practices as necessarily correct, (2) the replacement of rules of thumb by more carefully thought out guides to action, and (3) the collection of data to support decisions rather than reliance on casual judgments. Operations research draws on these principles and extends their power by engaging the computer's help. Mathematical models are used to characterize a process, object, or concept in precise mathematical terms. When models are incorporated into a computer program, it becomes possible to test new variables in order to ascertain how the process, object, or concept will behave under different conditions. A library administrator, for example, can thus learn how to avoid bad surprises by evaluating a course of action in

advance of its actual implementation. Johns Hopkins University and Purdue University are two institutions engaged in the application of operations research techniques to library management problems.

The computer industry, being more prolific than almost any other, announced only last year a third generation of computer hardware. Besides offering more computer power for the dollar, the new equipment has several engineering advantages. Its components are microminiaturized; many circuits are no bigger than a thumbtack. Its memory is large and can operate at speeds measured in billionths of seconds. To describe a billionth of a second, computer people have coined the word "nanosecond." A nanosecond is a very thin sliver of time indeed, if we consider that it is to a second what a second is to thirty years! The ability to process information at such lightning speed challenges man to be clever in the ways he chooses to use such power.

Another advantage of third-generation equipment is system versatility. It is possible to connect many different devices to the computer over standard communication links. With the availability of high-capacity, direct-access storage devices such as the Data Cell recently announced by IBM, we can now think in terms of storing large quantities of information at a central location, yet having the data available to numerous users at distant locations. This is what is meant by being "on-line" with a central computer and "time-sharing" its use.

The library network is the natural extension of the time-sharing idea, and automation planning at the Library of Congress best exemplifies the concept. It is based on the assumption that third-generation computers, when combined with communications equipment, will permit local and regional libraries throughout the country to have direct access to electronic stores of information in Washington.

The notion of direct dialing to a digital computer has already been tested at several experimental locations around the country. Elsewhere in this publication, M. M. Kessler of the M.I.T. Libraries describes the system he devised to link 100 users, over ordinary telephone lines, to an electronic store of information maintained by a central computer in Cambridge, Massachusetts. His interesting work is a microcosm of the kind of library network that may be developed in the future on a much greater scale.

The fact that there are no geographic limits to the library network idea is particularly attractive. Last summer, at an ASLIB Conference held at Ashridge College outside London, W. J. Bray, Deputy Director of Research of the British Government Post Office, described the gradual evolution of synchronous satellites and the effect they will ultimately have on world-wide communications. He traced the early history of telecommunications up to the most recent phase of laying cables by submarine across the Atlantic and Pacific

Oceans. By way of contrast, he then described how three synchronous communication satellites in circular equatorial orbit around the earth can reach any point on the globe, transmitting and receiving not only voice and video traffic but also the digital language of the computer. He predicted that communication satellites, such as EARLY BIRD, will permit intercontinental computer-to-computer "conversations," which eventually will facilitate the sharing of information resources among the libraries of the world.

The idea of creating a national library network strikes me as the most exciting development that data processing has thus far offered us. A library network is rich with the promise of a wholly new approach to the problem of gathering and retrieving essential information. It will transform libraries into active, rather than passive, sources of knowledge by permitting information exchange to flow in either direction between library and patron. The ability to confer with a library without necessarily visiting it, and the added ability to transfer selected information from library to home or office for individual use, should have a profound effect on the processes of research and education. If we, as professional librarians, are looking for an intellectual challenge, then this is it.

REFERENCES

1. "Hollerith, Herman." Dictionary of American Biography, ed. Harris E. Starr. New York, Scribner's, 1944. Vol. 21, Supplement 1, pp. 415-416.

Additional References

American Library Association. The Library and Information Networks of the Future. Prepared for the Rome Air Development Center at Griffiss Air Force Base, New York, RADC - TDR - 62-614, 1963.

Kessler, Maxwell Mirton. "The MIT Technical Information Project," Physics Today, 18:28-36, March 1965.

———. "The MIT Technical Information Project—System Description," Nov. 2, 1964, available as AD #608 502 from CFSTI, U. S. Department of Standards, Washington, D. C.

King, Gilbert William, et al. Automation and the Library of Congress, A Survey Sponsored by the Council on Library Resources. Washington, D. C., 1963.

Morelock, Molete, and Leimkuhler, F. F. "Library Operations Research and Systems Engineering Studies," College & Research Libraries, 25:501-503, Nov. 1964.

Ray, Robert, et al. Progress Report on an Operations Research and Systems Engineering Study of a University Library. (NSF-GN-31. Unpublished Report). Baltimore, Md., Johns Hopkins University, 1963.

THE TECHNICAL INFORMATION PROJECT OF THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY *

Project MAC at MIT has been mentioned as one of the newer developments in the computer art. In Project MAC, a man in Michigan can sit at a teletype machine and interrogate a computer, or in general behave as if the computer were next door to him. He can program, compute, or do what programmers call debugging or cleaning up a program. He can also, as of a few months ago, type a request such as "Compile a recent bibliography on laser physics" or "What's new in plasma physics?" It is this application of Project MAC that will be described here.

When the man in Michigan formulates his question, when he says "I want the latest bibliography on lasers," he does not get just a printout of a previously compiled bibliography. He gets the results of the actual putting together of a bibliography based on thousands and tens of thousands of documents in the computer's store, since it is a dynamic system. This then is a function that has been grafted on the MAC computer system at MIT, and it is necessary to understand the technological environment of Project MAC.

Project MAC was originally a very ambitious notion; some of the more venturesome people at MIT thought of it as "Machine Aided Cognition," which is a very revolutionary notion, but it is more realistically described as "Multiple Access Computer." TIP, the other acronym which will be used here, stands for Technical Information Program; it is that part of MAC which is concerned with non-computational uses of the computer, particularly for the manipulation, storage, and retrieval of all manner of technical information.

*The following paper is based on the tape of an unscripted talk given by M. M. Kessler (Associate Director of Libraries, and Director of the Technical Information Project at the Massachusetts Institute of Technology). Mr. Kessler has kindly granted permission for the paper to be printed in the version below, with the proviso that it represents the editor's understanding of his remarks; neither Mr. Kessler nor M.I.T. is to be held responsible for the precise form in which they appear here.

The essential idea of MAC is that a computer will be utilized as is a public utility. Just as there is not a little dynamo for each room or each building, so a computer for each department or each project is not needed. One central computer can serve the entire community. Now whether that community is the nation, or a region, or a university is not quite clear. It most certainly is not the nation, but it most certainly is not just one company either. At this time we might think of a "public utility" computer as being of the magnitude that might very well serve a region. Now obviously, in order for that to be realized there must be a remote access facility. And so the remote access idea is very much a part of the MAC thinking, and it means that you can approach the computer and use it, without any limitations, from a remote location. As of this time there are two possibilities. There is the telephone wire approach which is cheap and ubiquitous; all you have to do is attach a proper piece of equipment at the end of a telephone, which at this time is a standard teletype machine. A more sophisticated approach is to use a more broad band type of communication, such as television and data transfer circuits; but for any sort of human interaction with the computer, the telephone system is fast enough, accurate enough, and available enough so that we need think only in terms of telephone connections to the computer. And so the approach in and out of the computer, or the input - output machinery so to speak, is a standard teletype machine connected to the computer by means of telephone wire. Obviously remote access is a very important aspect of MAC.

The other important aspect of the public utility notion of a computer is time-sharing. What it means essentially is that people do not line up serially to use the computer, so that the tenth person has to wait until the ninth is finished, and the ninth has to wait until the eighth is finished, and so on. Instead there is a long trough, and they are all using it at the same time. MAC can accommodate thirty users at the same time; it can be time-shared by thirty people. Strictly speaking they do not time-share it, and they do use it in rotation. However the rotation is so fast that it is a matter of seconds, and the printout absorbs the cycle time because the computer works much faster than either a man can type or the teletype can type back to him. This means that, while one person is typing, if he hesitates at all or if his attention wanders, somebody else is using the computer. The internal machinery is such that the computer remembers the state of the last operation, and when it comes back to that operation, it picks up and continues another few milliseconds, and so on. It is not literally simultaneous, but virtually so. There are plans to increase this facility from thirty time-shared users to possibly one hundred and fifty.

Besides the notion of it being remote and of it being time-shared, there is the concept of being on line. And by "on line" is

meant that a dialogue can take place between the person and the machine. That is, you do not bring in the work, give it to an operator, and then come back the next day to find that one of your cards was misplaced. When you are on line, if you make a mistake the computer immediately tells you that it is a mistake, and if you want to correct something you can do it immediately. There is a back and forth real time interaction between the person and the computer.

A subtle psychological factor is involved here. When you work with a batch processing system that is not on line, and you bring some work in and are told to come back the next day, and you come back then and it is ready, you are very happy indeed that it is ready twenty-four hours later. But if you sit at a time-shared console and the answer comes back two minutes later instead of two or three seconds later, there is a great deal of impatience and annoyance. You don't mind waiting twenty-four hours, but you do mind waiting two minutes instead of two seconds. This is not entirely irrational because if you know that the answer is coming back in twenty-four hours you can do other things, but if you expect it in two seconds, and it comes back in two minutes, you just sit there and stew. This is an interesting human response, and there is apparently a critical time delay.

There are many other such things involved, of which this is an example, because we are dealing here perhaps for the first time with a real interactive process between man and an intelligent-like machine. At least it is an interaction between a man and a machine by way of man's intelligence. You are sometimes tempted to kick at your car or radio and take an anthropomorphic attitude towards them. This response is even more understandable where the machinery simulates some intelligence, and you expect more from it than it actually gives you; the psychological strain can sometimes be considerable when you realize that it is a poor dumb beast, and not even that but a dumb machine. This aspect of engineering or of research is extremely important in our type of application because we are thinking of using the computer in an intellectual manipulation and we really do not know what the answer to the problem will be. It looks pretty good if the machine can be made much bigger, much faster, and much cheaper. This is very important because if the machine is not cheap, you resent its not only wasting your time but wasting \$5 every minute or so, and that gets to be a very serious resentment. These are very important considerations, and they do color the use of the machine and its application in this area.

Fortunately, the computer people tell us that computer technology can only go one way; it will all become faster, cheaper, and smaller. It does seem that things are moving in this direction, so that we can look forward to these problems being solved. However from the beginning of computer technology (which is not more than

ten to fifteen years old) people have expected great things from the computer in the way of library and information application. As of today, if we froze the technology and took a snapshot of the situation, we must admit that the application or the contribution of computers to our type of problem has been rather trivial. Here and there a list is made up, or a little bookkeeping is done, or a circulation system is handled, but on the whole the computer has not yet made its impact on library and information sciences. However, the inception of the computer industry was only about twelve or fifteen years ago at the most, and the next ten or twelve years will really show a difference.

To proceed with the description of the MAC system, there is a computer, and the computer consists of two parts, the processing unit (with circuitry to do all the various computer manipulations), and the memory unit where all manner of information is stored. In the MAC computer, the memory is divided into several parts. There is first of all a part of the memory that has to do with the internal workings of the computer, e.g., the addition, subtraction, or whatever else you want to do with it; this concerns only the computer engineers, the maintenance people, and those whose responsibility it is to improve and develop the system. This is not visible to anybody; as far as the user is concerned, it is a black box that is closed, and he need not know what is inside. There is another part of the computer memory that contains within it various facilities which are available to the user. These involve the various computer languages, such as Fortran, MAD, or whichever it is, and this part may contain worked-out programs, so that if you want to perform a given operation you do not have to develop the program, you just call for it. This then is a library of publicly available routines, programs, and data.

The rest of the memory is divided into many compartments. Each compartment is accessible to only one user or subscriber, and out of each of these, schematically and not to dwell on the actual circuitry, come one hundred and fifty telephone connections, each going to a console, to a teletype machine. Thus there are one hundred and fifty locations where teletype machines are installed, and that location in Michigan is one of them. There are in fact about twenty or so scattered over the country; the rest are around MIT. Each user has at his disposal a common library of programs, languages, facilities, and so on. He also has at his disposal a private little library, his own office library, which he can put things into, take things out of, work on, and do all kinds of things that are available to him only. Within this range of memory, the strictest privacy is maintained. Great care has been taken so that user A can in no way get to user B's memory unless user B makes that facility available.

It is possible for any one of the users to take part of his private memory and declare it public, and that in every sense is a publication. As a matter of fact we have an editorial board and referees, and we

go through the entire procedure of publication before any user's private file is made available to others as a public facility. To give a routine example, suppose that somebody develops a new way of integrating an equation or a new way of solving a problem, and he thinks that it may be of use to others; he then submits it to the editorial board who examines it and approves of it for the public file. As soon as that is done, it is available to the entire community of users.

It is with regard to this sort of scheme that we have developed TIP which stands for Technical Information Program. A rather large slice of machine memory has been used for quite some time as our own private experimental slice, but about ten months ago it became sufficiently developed to be offered to the public. What used to be a private memory slice has now joined the public domain, and TIP is now available to all of the 150 consoles that have access to MAC. There are 150 consoles, but only 30 can use the computer simultaneously. In fact, since more than one person has access to any one of the 150 machines, there are some 500 people who at one time or another use Project MAC.

So much for the computer structure with which we work. Let us now look at TIP in more detail. When we first started thinking about this application, it was clear that we could not build an entire Library of Congress or MIT library system into a computer, turn the switch on, and then have it work. We had to have a model of the system, and the model had to be realistic in the following sense. First of all, it had to be scalable, so that if it worked, it could be scaled by a factor of 10 or 100; otherwise it would be a toy and not a model. The other requirement was that the model be big enough to be capable of functioning in a real environment and not be just an analytic type of model. If it was to be a model library, it had to be a big enough library so that people who want to use libraries would actually be motivated to come to this thing and use it. Even though they do not care about this as an area of research, they should care about this as a service. In other words the model had to have critical size. It had to be small enough so as not to tax the experimental facilities of the situation, and yet it had to be big enough not to be a toy but to be of serious interest to workers.

As a result, we had to limit our literature or holdings, and we picked physics as an area in which to work, more or less by chance because I am a physicist; but it was a happy choice because physics is a very well disciplined literature. More than that, we picked the journal literature of physics—no reports, no books, only physics journal articles. We process now twenty-seven journals, roughly 1200 articles per month. This corresponds to about 60 percent of the physics literature which ends up in Physics Abstracts. For each article in each journal as it comes in, we put into the computer memory the following information: what we call the identification, that is

the journal, volume, and page of the article; the title; the name of the author; the author's location or institutional connection; and whatever citations or bibliographical references are in the article either as footnotes or in the body of the text. There is no indexing or key word identification of any sort, and this is a calculated risk. One of the important considerations was that the system or model be scalable, and one of the most difficult things to scale upward is human intelligence. We wanted to see how far we could go with an input that is purely clerical. What we record for each article requires no judgment whatsoever, no assignment of key words, or indexing terms—it is purely clerical. Indeed this work is done by a girl who sits at the teletype machine (we do not use IBM cards) and types directly onto the computer memory the indicated information—identification, title, author, location, and bibliography.

The question of course is whether this is sufficient for a legitimate system. We went through a long series of experiments and we are satisfied now that it is sufficient, at least to begin with, and so we are not doing more. However the system is flexible, so that if at a later date we want to add other things, anything at all, we can do it by just typing in the identification data. The system is open ended, and if more information is needed, it can easily be put in.

This data then is on the computer memory disc in a format that is immediately available to any one of the 150 users. In other words, there is no loading of tapes or loading of cards; the information is there 24 hours a day, except during the times when the computer is off the air. It takes a bit of computer memory to do that but it is not prohibitive. This information is organized on the memory disc much as it would be organized on a library shelf. Think of twenty-seven journals, bound into volumes, and located all in one place; in our case we have files, and for this purpose each volume is a file. And within the file, within the volume, the information is organized by way of page numbers. This is the most primitive approach to file organization, and it is certainly not the optimal approach. It is now in the process of being changed to a more reasonable approach from the machine point of view because it is not the best approach for large scale searches. As our library gets larger, we will change this file structure, but as of now it is a serial file.

We then had to develop a set of words, a search language, because we set ourselves the design criterion that the people who want to use MAC for real purposes must not be asked to do programming. They are librarians, writers, working scientists—including physicists—but they are not programmers, and they are not to be asked to write their own programs. We had to develop a language that was close enough to English to make communication with the computer comfortable.

The language we have developed is a very comfortable sort of semi-English. What you do is to sit at the typewriter, log in, and identify yourself so that the MAC system accepts you as a legitimate user. You then type the word "TIP," which informs the computer that it is about to be used as a library and not as a computer. There are many other facilities which the computer has. For example, civil engineers have developed a road intersection program and bridge network program, and there are all kinds of biological and psychological programs; somehow the computer has to be informed of which aspect of its personality is about to be called into play, so to speak. When you type "TIP," that indicates that you are interested in the library part of the computer. The computer will come back and probably say "Ready," or some such thing; when it says "Wait," something is wrong. When it is ready, you might then say, "Search Physical Review, Volume 136 to 140," which is a typical search command. Or you might say, "Search Physical Review, latest issue," or "Search all latest," which means search the latest issue of all the twenty-seven journals. In other words, there is quite a variety of research statements you can make, each of which essentially means, "Take this designated literature off the shelf."

Then you type some request, like "Find title nuclear," which means "Find every title in this literature range that has the word nuclear in it." Here again there is a wide variety of possibilities. You can say "Find title nuclear energy," in other words you can use phrases, or you can say "Find nuclear energy," and in that case it will find every title which has the words "nuclear energy" regardless of their order, for example, "Production of energy in nuclear engines." If you want the words "nuclear energy" to be exactly as stated, you put an asterisk between them and this means that they must be joined. There is a wide variety of other manipulations; you can say "and/or but not," for example, "Find title nuclear but not spectroscopy." In other words, you want not the whole set of literature having to do with nuclear spectroscopy, but you do want other nuclear literature. You can say "Find nuclear and author such and such;" you can mix any of these things with the logical possibilities worked in. And of course you can find any one of these things separately. You can say "Find citation (and name the journal, volume, page)," and that means "Search this literature and find every article that cites this paper." Or you can say "Find author Smith," or any combination of these by way of "and/or but not," and so on.

By these directions you take the volume off the shelf, and look through it. Having found what you want, say all the articles with the word "nuclear" in their titles, what do you do with it? We have to make some sort of output statement, and a common form of output is printing. So you say "Print title, author, and page no." In other words, you direct the computer, having found all the papers with "nuclear"

in the title, to print the complete title, and the author, and whatever else you wish to ask for. You can also say "Save" because if you print, you erase what you have found as soon as it is printed out, but you may want to save the information for later work. Suppose you have discovered 230 articles with the word "nuclear" in their titles, you have created a new file, a new list. You have only to name it somehow if you wish to save it; so you say "Save file," and follow it by a name. For example, you might call it "Nuclear titles" or "List 1" or "Jones' favorite subject"; you must give it some name because you may want to come back to it later. Once you have done that, you can at another time say "Search" and instead of searching Physical Review, you can now search the file that you have made. This is a very important consideration.

The main point is that in this type of organization and language structure there is the possibility for searching by author, by any word in the title, or by what is known as citation index; all of these are available at any time and in many mixtures. You could do one, then the other, and so on and back and forth, and you can save that information and then look at it again later from another point of view. There are several safety features built in. For example, if you request an item that is not in the library, e.g., the Journal of Gestalt Psychology, the computer will say "The Journal of Gestalt Psychology is not in the TIP library." If you make mistakes in spelling and things of that sort, there are provisions for erasing and correcting them. If as you work, or as the computer works back at you, the telephone rings and you want to go away for a while, you can stop and then come back and type "Start," and it will start again where it left off, even if it is three hours later. We now have a graduate student working on a teaching program, so that the computer itself will teach people how to use the computer. This is of particular interest in a case where a non-programming population is involved. That is, we would like to provide for the situation in which a user logs in and says "I want to use TIP." He should know that much. Then the computer comes back and says, "Do you know how to do it or do you not?" If he says "No," a teaching program comes up point by point with illustrations. This, of course, is very closely related to the whole man-machine problem.

A second very useful search technique has been developed, something that we call bibliographic coupling. Let us say there are two or three articles which I know I am interested in. They may be my own papers, or they may be the papers of a friend who is an expert in lasers, and I know that I am interested in his work. Now I want to be able to conduct a literature search and find others like it. This can be done, provided that you identify the criterion of likeness. You may say to the computer "Find others like it," and by that you mean find other papers which share something with this paper. They may share a variety of things; they may share common words in the

title, or they may share the fact that they were all produced in or came from the same laboratory. These are useful but rather limited applications. What we found to be extremely useful is to say "Find other papers like these" where the criterion of likeness is the bibliography, or the list of citations in a given paper. Then we are saying in effect, "Examine these two papers that I know are of interest to me; observe the citations in these papers, then examine some given range of literature and find other papers that have similar items in their citations." We call this "bibliographic coupling," because the resultant papers are coupled not with words, but by virtue of sharing certain references. For example, if I write a paper with twenty items in my bibliography, and somebody else in Japan writes a paper with twenty items in his bibliography, and if it happens that ten are the same in both papers; then the probability is very high that these papers are related. We have done many hundreds of such experiments and find this to be an extremely powerful search tool. So the share group of programs, added to the find group of programs described previously, is very important. We can say "Share title" with the named paper, or "Share author" and so on; but of all those shared programs, the shared bibliography is the most potent search tool.

In the hierarchy of search procedures, we have the find list of programs and the share list of programs. Beyond that we have another even more sophisticated search procedure. (By the way, the number of papers on disc in the TIP library is now about 35,000 and growing at the rate of 1200 per month; this is pretty close to four years of literature.) Let us say we want to produce a bibliography on lasers. There are various strategies one can use. You can say "Search all," that means "Search the entire library," or you can say "Search 1965 papers" or anything of that sort.

Then you can say "Find title containing the word laser," and if you put a plus in front of laser, the computer will find anything where the suffix is laser, e.g., super-laser or anything of that sort; and sometimes it is convenient to do this. Then for output, you can say "Print and save," and call the saved file "Laser titles." In a typical case, this will result in about 230 papers. These are 230 papers out of those 35,000 which have the word laser in the title.

This by no means exhausts the laser literature obviously, because there are many papers that have to do with "lasers" which do not have the word laser in the title. But certainly those that do have "laser" in the title are important, and we have now compiled and saved this first list of all papers that have the word "laser" in their title. The next step is to say "Search this new list of laser titles"; we do not search the entire literature of 35,000 now, we search only 230 papers. Actually the direction is "Read laser titles," by which is meant that the computer is asked to list all the references in these papers in the order of their frequency of occurrence; out of 230 papers, there will

be perhaps 3000 references—slightly more than 10 per paper. Most of them of course, are referred to once and never again. Some are referred to twice, and so on; these 3000 references then will be listed in the order of their frequency of appearance in this group of papers, and we end up with a list that says “Paper 1 in this list of references appeared 17 times”; “Paper 2, 15 times,” and so on. To do all this will take 5 or 6 seconds.

Of course, one might take just a slice of this and ask for a list of all papers cited more than five times, and go look at those. But what is even more interesting is to say “Let us go back to the literature, search all of it again, and find all of those papers that do not have “laser” in the title, but nevertheless cite these same references.” For example, someone may have used a synonym for laser, namely optical maser, or some other term. This second complete search then gives you another sample of relevant articles. You look at this new set, and ask, “Here is a set of papers concerned with lasers, and the coupling is very good, although they do not have the word laser in their titles. What are the most frequent words in their titles?” In other words, “What are the most frequent and common actual substitutes for the word laser?” Perhaps you come out with two or three terms, such as optical maser or optical amplifier; then you go back through the literature, and search all titles for those words. This goes on back and forth, and each time you form a new list; there are in fact list merging programs that say “Go back to the original list, and if this new article is not there, add it; if it is a duplication, do not.” The list which started with 230 articles may add 50 in the next process, and 30 in the next; you continue searching as long as you wish, until you no longer add anything. Then you must be at the end.

This sort of procedure lends itself to many different kinds of manipulation, and we call them strategies. And this particular strategy just described, in which you start with a word and then go to frequency distribution of references, and back to the word and so on, is called Strategy A (since it was the first strategy which worked well for us). You can sit at the teletype machine now and say “Execute Strategy A on lasers;” you do not have to go through all of the steps just described, and the computer will proceed to perform each step in the whole process. This is a high order of programming which is extremely potent.

The system is now in use by something like 75 people per week. In other words, 75 people log in and ask for TIP to become available to them. We have written a monitor system, so that everybody who is using the system gets recorded as who he is, how long he is using it, what questions he asks and so on; and we are collecting user experience now through this monitor program. It is also being used by librarians in a much more self-conscious way. The system has been

in use about six months in this public fashion; we have had it in use experimentally for over a year now, but it was only about six months ago that it went public, so to speak.

THE JOHNS HOPKINS UNIVERSITY LIBRARY

Benjamin Courtright

The purpose of this paper is to describe two projects relevant to the interests of this conference which are active at the Johns Hopkins University. These are, first, the conversion of the shelf list of the University Library to machine-readable form, and, second, the design and operation of a new type of computer-based circulation system.

These activities were initiated and are being carried out at Johns Hopkins as part of an operations research and systems engineering study of university libraries, under the sponsorship of the National Science Foundation. The leadership in initiating and directing this study has from the beginning been that of Robert H. Roy, Dean of the School of Engineering Science and Chairman of the Department of Operations Research of the University.

The Project team has been fortunate throughout in having a close and cordial working relationship with the staff of the Johns Hopkins University Library. From providing a field for examination and case study of systems problems, to sympathetic and responsive reception of proposals for action, the Library has been friend, teacher, and subject all at once. This fortunate and indeed essential state of affairs has been due above all to the perception, the vision, and the understanding of the University Librarian, John Berthel.

The support of the National Science Foundation has been crucial. The cost of the work is small in terms of the long-range objectives which are at stake here, but without the far-sighted support of NSF these undertakings could not have been attempted at all.

Conversion of the Shelf List to Magnetic Tape

It is hardly necessary, at this date, to dwell on the desirability of having the shelf list of a library in machine-readable form. Suffice

Benjamin Courtright is a Research Associate in the Department of Operations Research of the Johns Hopkins University, Baltimore, Maryland.

it to say that Johns Hopkins arrived at this evaluation quite independently, in 1963, after a careful consideration of system problems. The orientation was definitely not that of electronic data processing (EDP), in the first place. In order to apply modern control techniques and optimization procedures to the problems of library organization and operation, we needed, in particular, data on the library's primary resource, its holdings. The shelf list is monumental, a single, unique file, ninety years in the making. For micro-examination, for looking at one or a hundred items in the collection, it was, and is, quite serviceable; but as a basis for control measures for the collection as a whole, for coping with the new generation of library problems, it was simply unusable. No one has ever regarded it as indigestible, only because no one has ever conceived of having to swallow it. Making this file assimilable became an integral, essential part of the solution to problems of inventory and flow in the library.

This account will omit a detailed description of the specific library problems which motivated the shelf list conversion, and the proposed solutions to them, as well as a consideration of the various means available to accomplish the conversion, and the factors affecting the choice to be made among them. Both topics have been covered in published progress reports. The following then is a case report of actual operating experience with the processing involved.

The conversion process decided upon was accomplished by the following steps:

- (1) Producing a working copy of the shelf list by microfilming.
- (2) Transcribing the desired elements of information for each record in the file and, at the same time, producing optically scannable copy, by typing the elements desired on a typewriter equipped with the special font required for the scanning process.
- (3) Automatically converting to magnetic tape on an optical scanning (character recognition) device.
- (4) Editing, reformatting, and updating the cumulated record on magnetic tape.

Microfilming—The need for a working copy of the shelf list is abundantly clear to everyone who has had contact with library procedures. The conversion process, including set-up time, took two to three months; to have attempted this from the shelf list itself, even if it had been possible to do it in the immediate vicinity of this unique file, would have produced endless confusion and frustration for everyone concerned. The choice of microfilming for producing the working copy has been a wholly satisfactory one. The cost of making the microfilm copy has turned out to be much less than the cost for any other available method. The completed microfilm copy is very easily

handled and stored; the microfilm replica of the entire shelf list fits into a box about 1-1/2 feet on each side. Above all, the microfilm copy did the job desired—it served as an adequate working basis for the conversion process. Although they had had no prior experience with the use of this medium for any job of this nature, we were delighted to find that the typists worked about as effectively from the microfilm reader as they do from original copy.

Microfilm does have one signal disadvantage: it is impossible to edit it or mark it for any intended use. Acceptance of this limitation forced the staff into the lengthy and laborious process of constructing a very careful and detailed set of decision rules for the typist to follow in reading and abstracting the data. Instead of depending on the editing of an experienced librarian or subject matter specialist the staff was forced into an explicit codification of the treatment for each of the many varieties of form encountered. The end product of such a process cannot be as good as that achieved through pre-editing by an experienced and competent person. On the other hand, the saving in time and in cost is very considerable; the process of forming explicit guidelines is in many ways a salutary one; and we believe that through the exercise of planned safeguards and checks, the final product will stand comparison very well.

One such safeguard, for example, was a direction to the typist to enter a special symbol "XXX" in place of the record element, whenever she encountered a form of entry which seemed to her not to be covered adequately by the rules. The last step in the conversion process has as one of its parts the printing of a list of records containing these symbols, by a simple computer program, for checking by a trained technical person. This means that such attention must be given only to the 2 or 3 percent of the records containing the anomalies, instead of to every record in the file, and was the last step in the conversion process.

The microfilming was done on a Recordak Model RM automatic microfilmer. This machine has an automatic feed, with a rated capacity of 375 cards per minute. The effective operating rate was about 200 cards per minute, allowing for the necessary shifting of cards and personnel breaks, and the entire operation was performed over one weekend, when the Technical Services Division was, of course, closed. The costs comprised \$480 for the rental of the microfilmer; \$229 for twenty-four 200-foot rolls of 16 mm. microfilm, including processing; and \$253 for the wages of the personnel employed for the purpose. A major part of the latter cost was for some sorting and re-sorting that had to be done on the file in the process. The total cost, then, was \$962.

Transcribing—Whether to convert all of the information on the shelf-list card, or only some part of it, was perhaps the hardest decision

of all. The question was considered from every point of view, over a range of possible alternatives. The final decision was made on the basis of the marginal utility of the various elements of information and on the basis of the specific functions for which the machine-readable record was proposed. A detailed consideration of these functions is set forth in the April 1963 Progress Report of the project. The fields of information entered into the record are as follows:

- (1) Call number. The complete call number is taken in every case. This includes the date of publication for modern cataloging.
- (2) Main entry. For corporate entries, the entire entry is transcribed. For personal authors, the last name and initials are taken. Authors' dates are omitted.
- (3) Title. This is the full title in most cases, a shortened form in the case of very long titles and certain other cases. Subtitles are omitted.
- (4) The number of pages, or the number of volumes in the case of a multi-volume set. (The number of pages for each volume is not given in the latter case, in local cataloging.)
- (5) The physical size of the book. This is, as usually given, the height in centimeters for normally shaped books, or both the height and the width where given under Library of Congress (LC) cataloging rules for unusually shaped books.
- (6) The number of copies of the work in the Library's collection.

These basic elements of information have been transcribed for every item in the Library, with appropriate coding so that any given field of the information can be retrieved at will. The records were transcribed, as indicated above, from microfilm readers onto typewriters equipped with the special font used for the optical character recognition process.

Optical Scanning—In a survey conducted to plan for the optical scanning phase of the process, two organizations in the country having fully operational pieces of equipment with sufficient sophistication to do the job were found. The company chosen was the Control Data Corporation (CDC), whose wholly-owned subsidiary, the Rabinow Engineering Corporation, had developed pioneer models of such devices. Having had no experience with either of the organizations contacted, the choice of CDC was made on quite explicit grounds; the scanning device used by the other company accepted a line of only seventy characters, in contrast to a line of 100 characters for the Rabinow machine; the operating rates of the other company's machine were much slower and the formatting less flexible than that provided by Control Data Corporation's operation. The location of the CDC machine, in Rockville, Maryland, has also proved

a considerable benefit, in view of the frequent coordination required in the planning and execution of the work.

This machine is a one-of-a-kind device, built specifically for in-house developmental testing and service bureau operation. Two different types of agreement were concluded with CDC, both of which have functioned satisfactorily. Under the first of these CDC took the microfilm copy of the shelf list and assumed responsibility for the conversion process. They subcontracted the typing work and assumed responsibility for the implementation of the conversion rules referred to above, subject of course to continuous coordination with Johns Hopkins in their application. Following the optical scanning process, they reprogrammed the raw output of the scanner, converting it into the bit code appropriate to the computers. The guaranteed accuracy for the whole operation, specified in the contract, is as follows:

(1) For the call number. Since this serves as an identifying code for the machine record, as well as for the many functions it has in its own right, the maximum obtainable accuracy was specified here. The terms of the guarantee are expressed as "virtually 100 per cent accuracy"; the procedure used was to make a visual check from the print-out of the final product against the original microfilm for this field, in addition to the error detection routines incorporated for the record as a whole, with both CDC and the typing subcontractor assuming responsibility for the checking.

(2) For the remainder of the record. The guarantee specifies that the number of errors shall not exceed 2 per 100 records.

These guarantees have seemed eminently satisfactory, probably exceeding the accuracy of most original copy. We are not as yet prepared to make a statement concerning our own sampling for verification of these levels, but there is no reason to believe, on the basis of experience, that CDC would fail to take full responsibility in the event that our checking shows these rates not to be maintained for some part of the completed record. Their responsibility would take the form either of cash rebates for our cost in instituting accurate correction procedures, or of their redoing the work in question.

The total cost for this processing is at the rate of 4.8 cents per record.

Under the other form of agreement, the typing is done by us, and the typescript is shipped directly to the scanner for automatic conversion. The output in this case is the raw output from the scanner written on magnetic tape; here we do the subsequent conversion processes and reformatting on our computer. The cost for this part of the operation is at the rate of 2 cents per line scanned. The scanning device accommodates 100 characters per line; our records exceed this length in 8 to 10 percent of the total cases.

The second form of agreement mentioned above was instituted in order to provide a solution to a specific operating problem, but the procedure has become interesting in its own right. The problem referred to is the conversion of items from a separate serials catalog. In the preliminary survey and planning, because the items in the serials catalog were sufficiently more complex in their bibliographic record than the monographs, it was preferable not to follow the first procedure outlined above in this case. After much exploring, a very intelligent and competent person with library experience was hired, to work solely on the serial records for a period of time. A suitable typewriter was obtained, an IBM Selectric, fitted with a typing element having the special scannable font of characters, and set up in a place immediately adjacent to the serial librarian's office for ease of reference. This was the first time that such a procedure had been tried, but after some initial difficulties it was found to work very well. Moreover, it was found that the overall cost, per record converted, is very nearly the same as under the other arrangement. Having no experience for a guide, it was expected that the unit cost might be significantly higher (although the total cost would in any case be small compared with the monographic conversion). The knowledge of this as an additional resource at about the same cost level seems highly propitious.

Total Cost—The total cost for the conversion of all records, including the microfilming, payments to CDC for their conversion processes, clerical wages, the purchase of magnetic tapes for the storage of the final product, an allowance for wages and computer time for the editing and reformatting processes, plus an allowance for overhead to the University, amounts to \$18,170. On a per character basis, this is at a rate of \$.0038 per word of 5 characters, or \$.00076 for each character converted.

New Circulation Control Procedure

The circulation control system outlined below was first proposed in a Progress Report issued in June 1964. Testing and development of the requisite means to put it into practice continued for several months after that; and in April 1965, shortly after the opening of the new central library building on the Johns Hopkins campus, the new system was put into operation for the Library as a whole.

Books brought to the circulation desk to be charged out are not opened. They are placed together on the counter, as many as are brought by a single borrower, and oriented so that the book labels

(bearing the call numbers) all face the same way. The borrower's card is set on a small stand facing in the same direction and the books and card together are then recorded on microfilm by a suitable camera positioned a few feet away (see Figure 1). Ten to twelve books can be photographed together in this manner, so that, with few exceptions, all of the books borrowed by one person are recorded at the same time. The charging process is completed by affixing a removable pressure sensitive label, imprinted with the due date, on the outside back cover of each book (see Figure 2). The idea of the labels was originated in connection with the system in order to obviate altogether the need for opening books, so as to speed up the charge-out process, and to provide a more easily noted reminder of the due date.

At convenient intervals the film is developed in a processing unit and then placed on a microfilm reader situated alongside a key-punch machine. The reader currently in use here projects the microfilm image at a magnification which is slightly greater than the reduction ratio of the camera lens, so that the image of the books is somewhat larger than life-size. The call number only is read from the book label and key-punched, together with the borrower's card number by the operator (see Figure 3). Experience has shown that if the call number can be read on the book itself, it can be read from the projected microfilm image. Since this number has had to be legible in order for the book to be in place on the shelf, it is generally read without difficulty; in the occasional case of a torn or mutilated label, however, identification can also be made from author-title information appearing on the cover of the book, the call number then being obtained from the Library catalog. This does not occur frequently.

The discharging of returned books follows the same flow described above. The borrower drops books off at the circulation desk in the usual manner, and the attendant simply photographs them with a card marked "Discharge," in place of a borrower's card (see Figure 4). The key-punch operator converts the call number as above, putting a one-character symbol denoting "discharge" in the place of a borrower's number. Renewals are handled in an analogous manner.

Capabilities of the System—Given the charge and discharge records on punch cards, all of the remaining functions of circulation control are accomplished by a computer and computer program. The program was originated by Dr. Willis Gore, Associate Professor of Electrical Engineering of this University. It performs the common tasks which are always associated with circulation, and in addition has the ability to do a number of things which seem highly desirable but which have rarely if ever been feasible in a large library.

In the category of common tasks, the computer, guided by the program, keeps track of all outstanding charge-outs and deletes the record for items returned; it determines when an item is overdue and prints out notices bearing the proper address, ready for mailing; it calculates fines, where appropriate, for items returned late; and it provides a readily accessible record of the whereabouts of all materials in circulation, or in use internally within the Library.

The capabilities of the system which are new in some sense do not form a well-defined list, but they include:

(1) The ability to compile circulation statistics for each item in the collection and, equally important, to present this information in concise, usable form, for any selected set of material.

(2) The ability to characterize circulation reliably in terms of various categories of borrowers, the types and subject areas of materials borrowed, periodic changes in rates of circulation and long-term trends, or any combination of the foregoing.

(3) The ability to print out at any time a listing of items charged out by individual borrowers.

None of the capabilities listed above is new in the sense that its attainment is theoretically impossible under older systems of circulation control, but in general the work involved would be tedious and time-consuming. Under a manual card file system or any of its logical derivatives, the file can, in the nature of the case, be kept in only one order. Conventionally this is an arrangement by call number. The file however has many other attributes, or dimensions, of potential interest: borrower, category of borrower, type of material, main entry, time and date of borrowing, etc. To retrieve information for any of these other dimensions, the librarian must either make a painstaking search through the whole file, employ a system of mechanically tagging or reordering throughout the file, or generate and maintain a duplicate file for each different approach.

In a computer-based system, on the other hand, the file may be regarded for practical purposes as multi-dimensional; roughly speaking, using it in any dimension does not affect its integrity in any other. There is, of course, a cost associated with the computer's operations, but the cost is small relative to the value which will be obtained for the capabilities listed above and, with the exceedingly rapid growth in the technology of computers, it seems likely that this will be true for other functions in the not too distant future. In addition, the "real time" taken by the computer operation is negligible, in marked contrast to the length of time required for such purposes under manual systems.

Means of Implementation—Encouraged by the Library staff, we embarked upon the conversion of all circulation control processes in the

new Johns Hopkins Library building, then in the final stages of construction, to the proposed design. Our lack of trepidation was matched only by our lack of knowledge of the problems to be encountered. The microfilming of the shelf list was a considerable task but that was a straightforward application using a standard piece of equipment, quite a different affair from finding, modifying, and fitting together pieces of equipment to perform new functions. We have since learned a small amount about lenses and prisms, microfilming equipment, films and film processing, pressure-sensitive label materials, etc., but in the main it has been the humanity and forbearance of our associates and the Library staff which have made the development possible. Our boldest claim is that, if the system works with what we have put together here, it should work anywhere.

The equipment installed at the circulation counter is in a somewhat different configuration from that originally planned. The original design contemplated a camera set in an overhead position, shooting down on the books and borrower's card, with the books placed spine upward in the usual case. The architectural design for the circulation area in the new Library building had, however, already been drawn up at that time, and it turned out that a configuration such as that proposed would be artistically and architecturally unacceptable. Accordingly, the photographic recording process was adapted in such a way that the books and the borrower's card would be photographed on the horizontal axis. The books, therefore, are set on the counter on their ends, in the position they normally occupy on a shelf, with a suitable movable block to hold them in this position, while the transaction is recorded.

In order to avoid having the camera on top, however, thus cluttering up the new circulation counter, it was placed just to the side and below the counter top with a reflecting surface set at a 45-degree angle to direct the image down into the lens (see Figure 5). At first a front surfaced mirror was used for this purpose, but a total-internal-reflection prism does the optical job at least as well and is a great deal easier to support in a proper position and to keep clean.

In operation the camera is completely enclosed in a housing constructed for the purpose, with a removable side to allow for the insertion and removal of film. The housing serves three purposes: it keeps the camera anchored firmly in position, since the housing is bolted securely to the supporting shelf and the camera is fitted to its interior; it provides physical protection to the camera; and it is lined with soundproofing material, to reduce the slight noise of the relays in the camera during operation. The camera itself is an automatic microfilm unit removed from its normal supports and slightly adapted for the purpose. It is actuated by a toggle switch which has been connected to the operating relays and located remotely, convenient to the location where books are placed for charging. The toggle switch is

spring loaded, and simply pushing it down takes the exposure and automatically advances the film ready for the next transaction. It is an exceedingly reliable piece of equipment; in addition to a highly developed engineering design, it has a comprehensive alarm system which gives a signal whenever the camera is not in condition to take the exposure properly—if the supply reel of film is nearly exhausted, if the film is not properly seated in its channels when reloaded, or in almost any other anomalous situation which might occur.

The film and pressure-sensitive labels now in use are still being explored. The materials in use do reasonably well, but work is still very much in progress to improve their efficiency and performance. The labels are secured with a background design preprinted on them; the due date is imprinted on them by a small automatic label imprinter secured for the purpose, and they are available at the circulation counter by an automatic dispenser which feeds out a new label each time a label is removed. The cost is less than a tenth of a cent per label at the present time and should be reduced with further development.

All computer operations have been carried out on an IBM 1401. Although there are two IBM 7094's in the University's main computing facility, the smaller machine has seemed preferable for this work. It is located right in the new Library building, so that access has been excellent; and it is more suited in basic design to the type of processing being done. This may be subject to drastic change soon; the University has filed a "Letter of Intent" for an IBM 360 Model 50, to replace the IBM 1401. The decision is not yet firm, but the prospect is an exciting one; if it is found that the circulation control operation can afford an input-output console at the circulation desk, and if a suitable time-sharing package can be included in the computer's operating system, a true on-line circulation system would seem to be a possibility.

The IBM 1401 in use has 8000 positions of core storage, four 729 magnetic tape drives (no disk storage), and a fairly complete array of special features on the central processing unit. The 8K storage limitation is a decided nuisance, but so far it has been possible if somewhat awkward to program around it. The daily circulation control update program, for example, has nine overlays.

Programming has been done entirely in the Autocoder language. In addition to the basic circulation control package, there has been the task of assembling data for and issuing library user cards; prior to this, library cards were not issued at the University at all. Identification numbers for circulation purposes are generated in the computer, and the cards are prepared on preprinted continuous forms on the computer's printer, for the several classes of users of the Library.

The amount of time required for making indicated modifications and changes in the programs has been rather surprising. The original circulation control updating program was carefully made, and yet, as experience has been gained, the need for adaptations and the opportunities to make improvements have been virtually continuous. By far the largest part of the author's time for the past several months has been devoted to writing programs, debugging, and operating the computer. On the basis of this experience, it seems that the improvement of means for man-machine communication remains a most important objective.

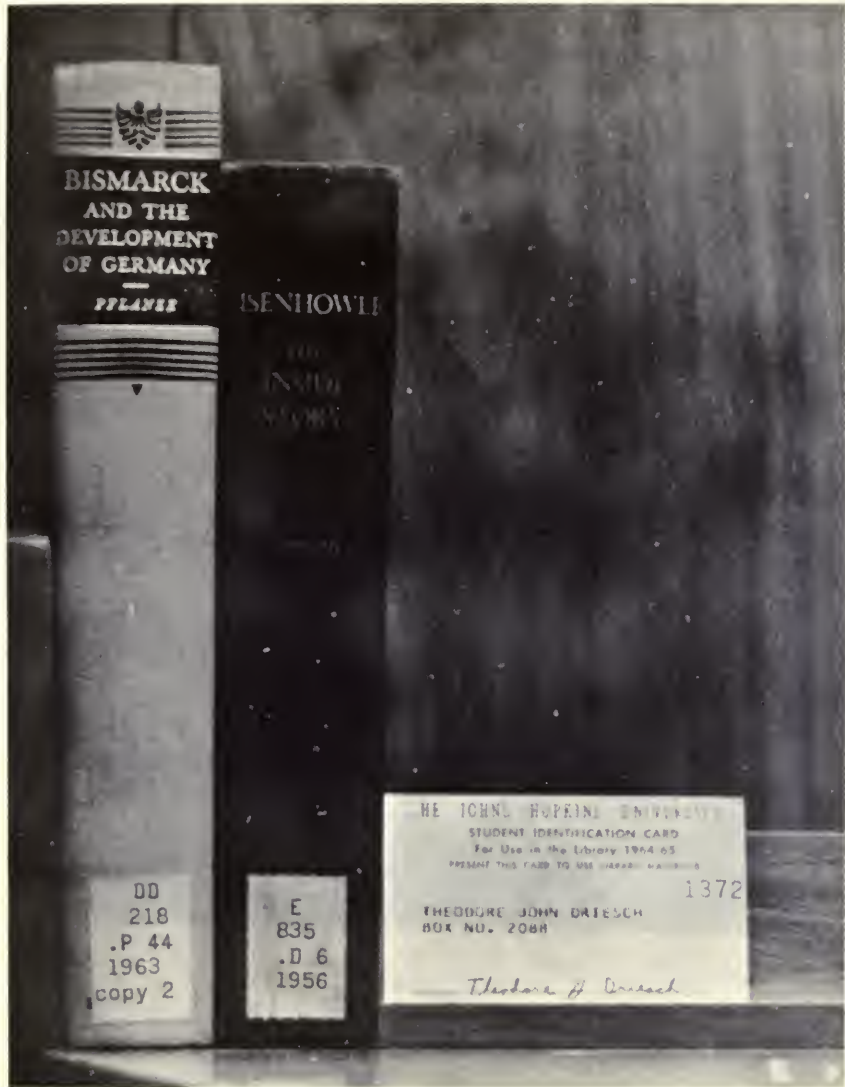


Figure 1
A Frame of the Microfilm Circulation Control Record



Figure 2
Pressure-Sensitive Label with Due Date



Figure 3
Microfilm Reader and Card Punch



Figure 4
Returned Items in Position for Discharge



Figure 5
Close-up of Camera

THE USES OF PLATO: A COMPUTER CONTROLLED TEACHING SYSTEM*

Donald L. Bitzer, Elisabeth R. Lyman, and John A. Easley, Jr.

The use of a high-speed digital computer as a central control element provides great flexibility in an automatic teaching system. Using a computer-based system like PLATO permits versatility in teaching logics, since changing the type of teacher merely requires changing the computer program but not the hardware. In addition, having access to the decision-making capacity of a large computer located as one unit permits complicated decisions to be made for each student. Such capacity would be prohibitively expensive to provide by means of decision-making equipment located at each student station. Studies of queuing that occurs with multiple student requests show that the system could teach as many as a thousand students simultaneously without incurring a noticeable delay in processing any student's request.

The educational results thus far have been extremely encouraging. However, reliable conclusions on educational achievement must await the results of more thorough experiments now in progress which

*This project was supported in part by the Joint Services Electronics Programs (U. S. Army, U. S. Navy, and U. S. Air Force) under Contract No. DA 28 043 AMC 00073(E) and by the Advanced Research Projects Agency through the Office of Naval Research under Contract No. Nonr-3985(08). Portions of the material in this report have appeared in Computer Augmentation of Human Reasoning. (Symposium). Washington, D. C., Spartan Books, Inc. 1965. Chapter 6, pp. 89-103.

Reprinted with permission from January 1966 Issue of Audio-Visual Instruction.

Donald L. Bitzer is Research Associate Professor and Elisabeth R. Lyman is Research Assistant Professor in the Coordinated Science Laboratory, University of Illinois, Urbana; John A. Easley, Jr. is Associate Professor of Education, University of Illinois, Urbana.

include larger numbers of students learning under a variety of conditions. The adaptability and usability of the system for a variety of purposes in education (including the behavioral and physical sciences) have been clearly demonstrated.

Introduction

During the past five years, the Coordinated Science Laboratory at the University of Illinois has developed and experimented with an automatic teaching system called PLATO in order to explore the possibilities of automation in individual instruction. The PLATO system utilizes a high-speed digital computer as the central control element for teaching a number of students simultaneously, while still allowing each student to proceed independently through the lesson material.

Three successive models of PLATO have evolved, each embodying improvements indicated by the previous model. The first consisted of a single student station connected to ILLIAC, a medium-speed computer built at the University of Illinois.⁴ The second model had two student stations, was connected first to ILLIAC and then to a CDC 1604 computer, and was used to study the problems created by multiple student use of the system.¹⁰ The third and current model has 20 student stations connected to the CDC 1604 computer.

The rules governing the teaching process are included in the program read into the central computer. A complete set of rules is referred to as a "teaching logic." The Coordinated Science Laboratory has experimented with two basically different types of teaching logics, a "tutorial" logic and an "inquiry" logic. A tutorial logic is designed to lead the student through a fixed sequence of topics, but it also provides branching between problems (which is under the student's control, voluntary or involuntary). In a lesson that uses the tutorial teaching logic, the system first presents facts and examples and then asks questions covering the material presented. The student composes answers, and when he is ready, he asks the system for a judgment. If he finds the questions too difficult, he may branch to easier material. Involuntary branching occurs when evaluations of the student performance are included in the lesson program, which prescribes branching if predetermined criteria are met by the student.

An inquiry teaching logic, on the other hand, can be characterized as a system permitting dialogues between the student and the computer. Typically, in a lesson that uses an inquiry teaching logic, general problems are presented to the student. To solve them, he must request and organize appropriate information from the

computer. In such a teaching logic, the student may be asked to demonstrate his achievement by answering questions, but he may also seek information within a given range of possibilities in order to answer such questions.

Both types of teaching logics and a variety of lesson materials have been employed in exploratory studies in order to test the capabilities of the system. Some of these exploratory studies have investigated system variables such as data rates between the students and the system.² Other studies have dealt with the psychological aspects of the lessons and variations in the teaching logics.^{1,7,8}

The PLATO Teaching System

Student Stations. A block diagram of a single student station in the PLATO teaching system is shown in Figure 1. The system provides for communication in two directions. Each student is provided with both an electronic keyset as a means of communicating with the central computer and a television screen for viewing information selected by the computer. The student's main keyset resembles a typewriter keyboard, and the keys can be assigned any functions the teacher desires. Usually, the alphanumeric characters are assigned positions similar to those on a standard typewriter keyboard, and punctuation, special characters, or special control functions are assigned to the extra keys.

Electronic Book. There are two sources of information which are usually displayed on the student's television screen. These sources (called an electronic book and an electronic blackboard) are diagrammatically shown in Figure 1. The electronic book consists of a bank of slides prestored in an electronic slide selector which is controlled by the computer. In the latest model of PLATO, the random-access slide selector stores 122 slides and has a slide access time of less than a microsecond. Information stored in the slide selector is the type that would usually be found in a textbook or in class notes. Although the slide selector is shared by all of the students, the students can view the same or different slides simultaneously. This is accomplished by having the video information available from all slides concurrently, and by connecting electronically the student's television display to the proper video output.

Electronic Blackboard. The electronic blackboard consists of a computer-controlled storage tube for each student station. Diagrams, symbols, and words are plotted in a point-by-point fashion on the student's storage tube. Approximately 40 alphanumeric characters can be written on the student's blackboard per second, and the entire blackboard can be erased in two tenths of a second. This

arrangement permits information (that cannot be predetermined) to be presented to the student, such as information generated while teaching the student. For example, the system can display a sketch of an experiment the student has requested or an answer the student has composed which cannot possibly be anticipated. The images from the blackboard and the electronic slide selector are superimposed on the student's television display, enabling the student to fill in blanks on the slide and compare his answer with the question. Figure 2 shows a block diagram of two student stations, indicating the shared and the individual parts of the system. Information for a student can appear on his television screen from either the blackboard or the book or from both simultaneously.

Teaching Logics for the PLATO System

Original Tutorial Logic. The tutorial logic was the first of the two main types of teaching logics explored on the PLATO system. In this teaching logic, the keys were divided into two types—those used for inserting constructed responses to questions and those used by the student to control his progress through the lesson material. The lesson material was organized into two types of sequences: the main sequence, consisting of the minimum material that must be used by all the students, and the help sequences, provided for students who had difficulty with questions in the main sequence.

The student began by viewing text material in the main sequence. When he completed reading a page of text, he proceeded to the next page by pushing the button labeled "continue," or he returned to a preceding page by pushing the button labeled "reverse." As the student proceeded through the lesson, he was presented with questions. The teaching logic required that all the questions on a page be answered correctly before the student could continue. The student was allowed as many attempts as necessary to answer the question correctly. If he had difficulty with a question, he could push the button labeled "help," which took him into a help sequence pertaining to the question. After completing a help sequence, the student automatically returned to the question he was trying to answer in the main sequence.

In a later version of the original tutorial logic, as developed by Braunfeld, different types of wrong answers called for different help sequences. An error detector was used for automatically controlled branching. In addition, the later version of the original logic permitted some questions to be designated for monitoring by an evaluator in the computer program. The student's responses to monitored problems were used to determine whether he was branched forward to the

next section of the main sequence or routed through material designed especially for students who failed the criterion test in the evaluator.

In order to prepare lesson material for the original tutorial logic, one had to organize the material into a set of slides (with at least one help slide for each question in the main sequence) as well as prepare a parameter tape. The parameter tape contained the answers to the questions, their location on the slide page, and the order in which the slides were logically connected. If the special help sequences and the evaluator were used, error categories had to be specified for the error detector and a list made of monitored problems and their criteria for evaluation.

The most recent version of the PLATO tutorial logic, which is much more generalized than its predecessors, will be described in a later section of this article.

Inquiry Logic. While the tutorial logic serves well for many purposes, there are types of problems in which even more control should be given to the student as well as an opportunity to ask questions of the computer. To accomplish this, the inquiry teaching logics were written.

An inquiry teaching logic permits a student to request information. The computer correctly interprets the request and replies from stored information or calculated results. This logic provides, in effect, a syntax for the student to use in communicating with the computer. The student directs his learning by composing his own requests.

In the tutorial logics, the student communicates with the computer either through one of the control requests—turn the page, judge my answer, give me help—or he composes short answers which usually must match one of the several alternative stored responses. If he should type a question such as “What does ‘exponent’ mean?” the computer would only respond with a “no” since it treats the student’s response as an answer. However, the inquiry logics provide a syntax by which a student can ask questions about the lesson he is studying. The syntax he uses can be viewed as a tree of choice points in which selections are made at each choice point.

Figure 3 presents a simplified flow diagram of a simulated laboratory, illustrating the general form of classification syntax. By pushing the button labeled “lab,” the student is shown the general categories of available information. Having chosen one of these categories, he is shown more detailed selections within that category. In general, successive subcategories can be chosen until the detailed classification is specified. However, it is often desirable to have the major categories specified independently, e.g., object, conditions it is exposed to, and particular properties about which information is desired. In such a case, the student can pass repeatedly through

several successive levels of selection, once through for each general category. Specifications made within one general category can be stored and used in conjunction with those made within another category. When the requested information has been completely specified, it is displayed on the student's television screen.

Many variations on this classification scheme are possible. Figure 4 shows how a student might have set up two experiments in a simulated laboratory in which the property to be measured is chosen first and the object and the condition specified later. The properties about which information can be obtained are the weight and overflow volume of objects listed. The conditions available are the liquids in which an object is immersed. Figure 4 also illustrates the use of both graphical and numerical display of results. Figure 3 shows only two choices at each choice point, and Figure 4 shows two for the first choice and five for each of the next two in the volume experiment (and six and seven respectively for each of the next two in the weight experiment). Figure 3 illustrates only four specifications on each pass through the tree, but two passes provide 16 combinations of objects and properties. Figure 4 provides 50 possible specifications on one pass through the volume experiment and 84 through the weight experiment. Some combinations have been used involving two passes. Ten choices at each choice point on three levels would permit a thousand specifications on each pass through the tree. PLATO III permits 128 selections at each of 128 choice points, which should be more than adequate for any foreseeable educational purposes.

Specification at choice points may seem a somewhat artificial way of asking a question, but it resembles the way one locates merchandise in a department store, and even elementary school children adapt to it easily. It requires only a slight rearrangement of ordinary language. For example, instead of typing "What's the effect of administering nitroglycerine on the heart rate of the patient?" the student in a PLATO teaching program for nurses who wished to ask this question typed coded numbers for the following sequence of phrases: return patient to original state, give drugs, select nitroglycerine, check condition of patient, vital signs, pulse rate. At this point the computer answered with the pulse rate.⁶ Students quickly learned the syntax required and usually formed such coded questions more rapidly than they could type them in English. The computer responded immediately, displaying information obtained by computation or from memory. The student proceeded to try other experiments until she was confident concerning the treatment of the patient.

An inquiry logic written for the PLATO system which deserves special comment is one that permits the student to solve mathematical problems that require many lines of work and for which all possible solutions cannot be anticipated.⁹ In this teaching logic, the student is informed whenever he violates any of the rules of

mathematical logic. The computer does not store a set of correct solutions, but it does store the mathematical principles available to the student. The rules of mathematical logic are built into this teaching logic by means of decision programs. Thus, this logic simulates a teacher who watches students at work and tells them whenever they make an error but doesn't tell them what they should have written. The student is, in effect, asking whether each move he proposes is a valid one, a question to which he gets an immediate reply.

The PLATO Compiler. A PLATO compiler was developed in 1964 which permits simple preparation of all types of new teaching logics. With this compiler, educational researchers have prepared several new teaching logics suited to their own purposes in fields ranging from mathematics to the behavioral sciences.

Preparing an inquiry type teaching logic requires specifying the tree structure of the syntax the student uses to communicate with the computer. Preparing a tutorial logic also requires specifying the structure which the student or teacher uses in communication decisions with the computer. The PLATO compiler permits the logic designer to specify for each choice the next choice point to which each response leads. Each choice point can present a slide, some message printed on the blackboard, operate a piece of auxiliary equipment, etc. All of these details are specified in pseudo-English. Special decision rules are written as necessary using an augmented Fortran language.

All of the PLATO programs or lessons written since the fall of 1964 have been written for the compiler. Many of the old lessons have been revised and reprogramed using the compiler.

New PLATO Tutorial Logic. The new PLATO tutorial logic, written for the compiler, allows very flexible rules for the teacher.³ The teacher may allow the student to respond with long answers. Several help sequences are permitted, and many judgers are available, including a spelling judger (which prints "SP" instead of "NO" on the blackboard when a spelling mistake is made). Eight special effects are available for 16 different keys, such as disallowing certain keys at specific times in the lesson or introducing an inquiry procedure such as curve plotting, available upon student request. Special remedial or challenge sequences are possible. A comment page allows a student to make comments on the lesson at any time; an instructor page allows the student to communicate with the instructor via the PLATO display. Finally, as the most important feature, the new logic contains an author mode so that the teacher may insert or change page answers and page descriptions on line with the computer.

Interconnection of Student Stations. Although independence of student stations was initially thought desirable, many uses of station interconnections were later suggested. The interconnection was accomplished with a short addition to the resident computer program.

This development has allowed teacher-student interactions, negotiation studies, and concept development exercises.

New Logic for Problem Solving. The more generalized version of the mathematical problem-solving logic is being written with the use of the compiler. This logic, incorporating improvements indicated through experience with the prototype, will allow the student to formulate his own problems and conjectures and work them out with the same supervision as if they had been problems stored by the author of the lesson. The judgment of student errors could be postponed, if desired, until the student requests that his work be marked. It is expected that this logic will be able to cope with problems in elementary algebra, logic and set theory, and some portions of geometry.

Student Records. One of the important features of the PLATO system is the "perfect workbook" of student performance kept by the computer. The student records include a record of each button pushed and the time at which it was pushed. This information is available in two forms: one form is a printed history of events that can be immediately scanned by the teacher; another form is one stored on magnetic tape that can be processed by the computer for a detailed statistical analysis.

Exploratory Studies Using the PLATO System

Student Performance and Queuing Studies. Several studies, some of which have already been mentioned, have been completed using both the tutorial and the inquiry teaching logics. Lesson material drawn from mathematics, computer programming, and electrical engineering initially were programmed with the original tutorial logic. Most of these studies employed approximately 10 to 12 students as subjects, each of whom attended three of four one-hour sessions. Results of some of these studies are available in another report.²

Briefly, the results from the early investigations showed the following: (a) There was no significant difference between the post-test scores of students who received instruction via the PLATO system and those who attended regular class. However, the amount of time spent on the lesson material was significantly less for the students working on PLATO; and (b) using over 50,000 student requests obtained with the lesson material, queuing studies were performed. It was determined that a general purpose computer, having a high-speed capacity of 1.5 million bits, would allow 1,000 students to be tutored concurrently on eight different lessons without incurring a noticeable delay on any student's request.

University Courses. Recently, the new PLATO tutorial logic was used to program half of the material for a semester's work in a course in circuit analysis offered to electrical engineering junior and senior students. Although no detailed analysis or evaluation of the students' responses was made, some of the more obvious results showed that the students appreciated flexibility in the system, enjoyed features such as curve plotting, and thought the course material was markedly clarified by the PLATO lessons. At present, the logic is being used for credit courses in "How to Use the Library" and "Fortran Programming for Business and Commerce Students" as well as for the electrical engineering course. Evaluation of student performance will be made from the detailed records provided from the system.

Text Testing. Worthy of mention is a study now in progress which uses a logic basically tutorial in nature to record performance of students as they test new textbooks. The student works freely through a textbook, which is reproduced on the PLATO system, answering problems or questions at will. The on-line author input allows on-the-spot changes and revisions by the author. Data retrieval programs will give the author a variety of information useful in his next revision.

Studies Using Auxiliary Equipment. It should be noted that the PLATO system can include auxiliary devices operated under computer control. The inquiry training lesson used a computer-controlled motion picture projector.⁵ Physiological recording devices have also been used with the system.¹ A more unusual study is one substituting a piece of experimental apparatus for a student at a station, with input from the experimental setup replacing the operator response at that station. A student at a second station can manipulate a real experiment through his station without ever touching the apparatus and can obtain the experimental results on his display.

Teaching with the PLATO system can be extremely varied since laboratory as well as classroom work is possible. Experiments may be performed which are either real (like those just described) or wholly simulated (like those referred to in the discussion of inquiry logics).

Other Research. Other teaching research projects have included drill sequences for remedial arithmetic studies, physiological studies relating to mathematical discovery,¹ and work in the area of verbal learning and retention. The wide range of exploratory studies possible with the PLATO system serves to demonstrate the versatility and flexibility of the system.¹¹

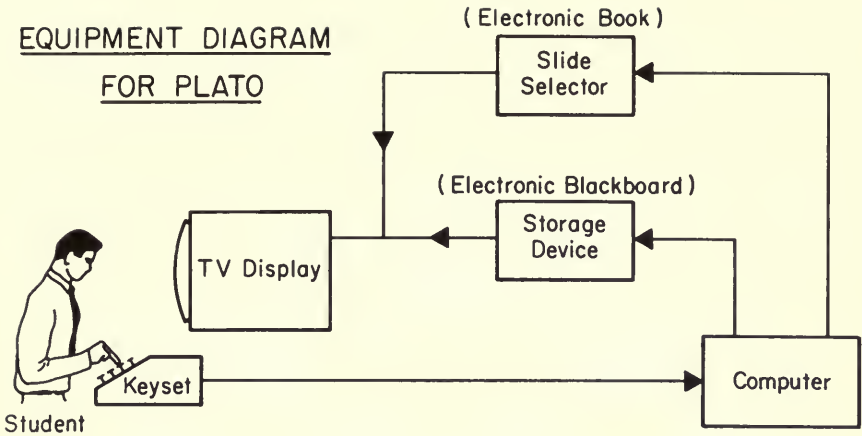


Figure 1
Block Diagram of the PLATO Teaching System (One Student)

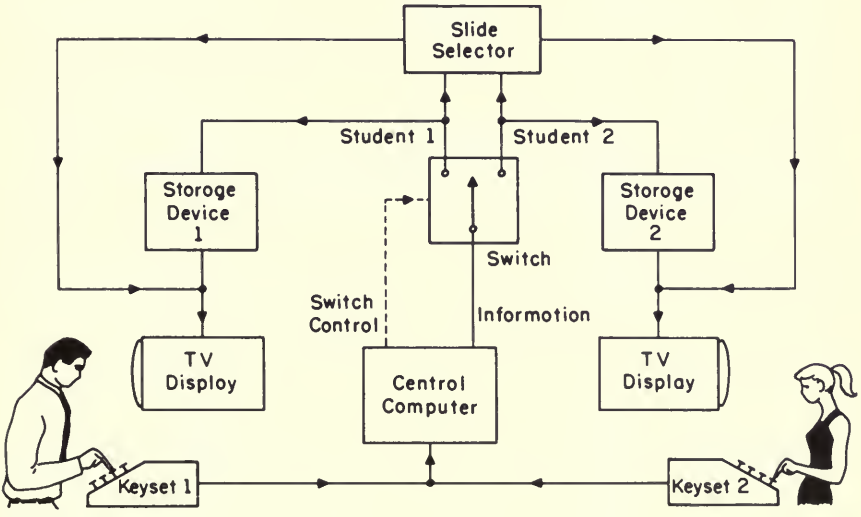


Figure 2
Block Diagram of the PLATO Teaching System Showing Shared and Individual Parts of the System

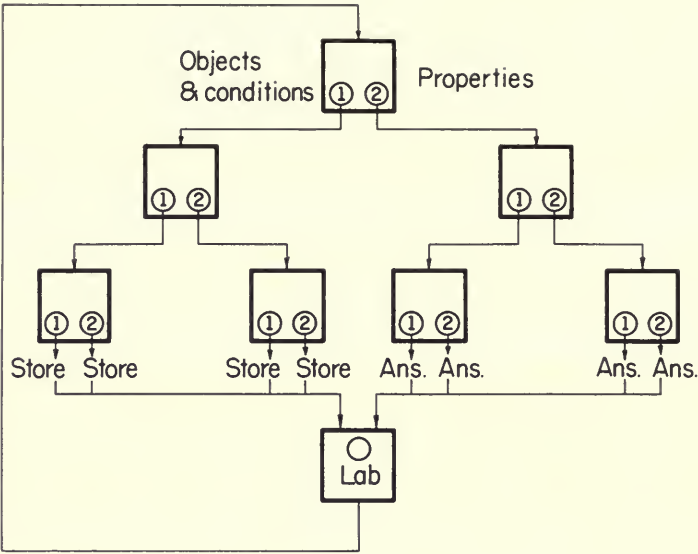


Figure 3
Simplified Flow Diagram for Simulated Laboratory Experiments
Using PLATO III Inquiry Logic

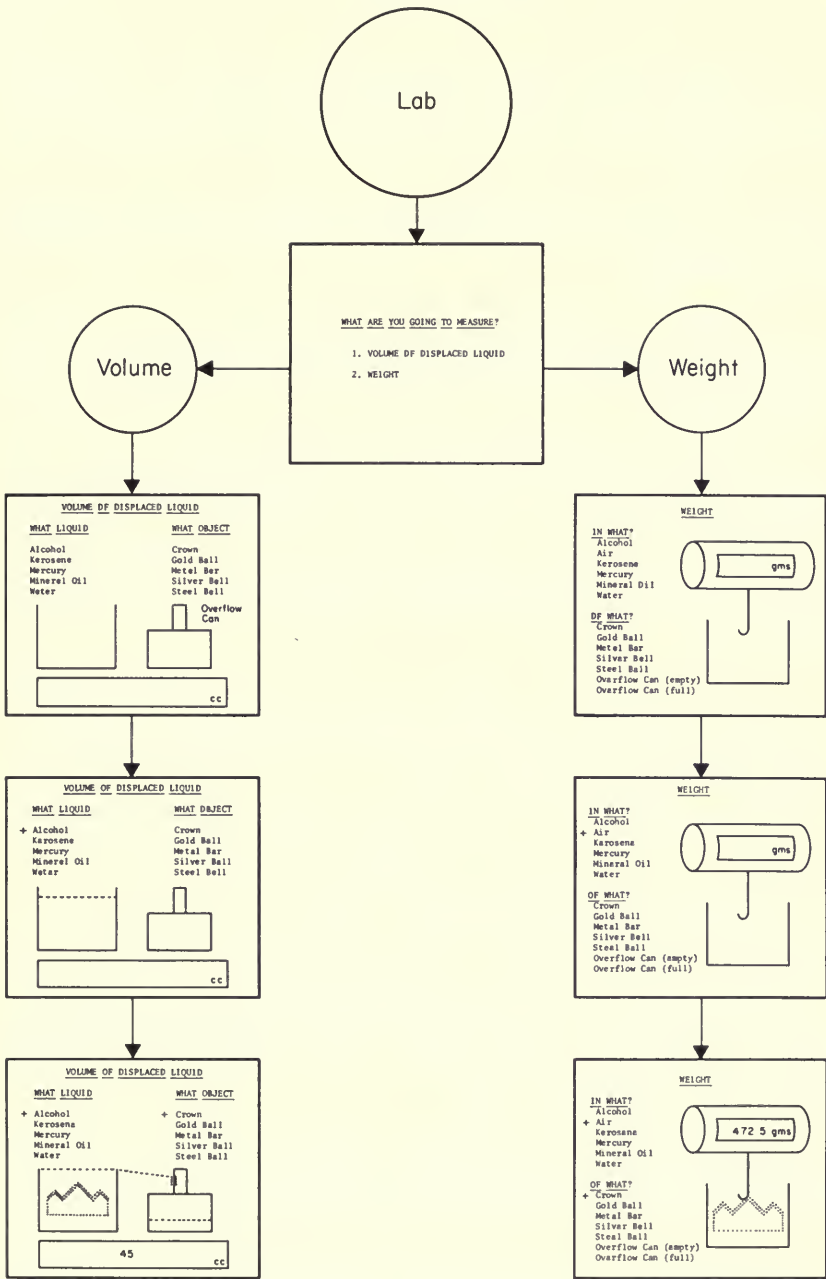


Figure 4
Example of Student's Use of an Inquiry Teaching Logic

REFERENCES

1. Avner, R. A. "Heart Rate Correlates of Insight." CSL Report R-198. Urbana: Coordinated Science Laboratory, University of Illinois, 1964.
2. Bitzer, D., and Braunfeld, P. G. "Description and Use of a Computer-Controlled Teaching System." Proceedings of the National Electronics Conference, Vol. 18. Chicago: Rogers Printing Co., 1962. pp. 787-92.
3. Bitzer, D.; Chan, S.; Johnson, R.; and Walker, M. "Lesson Preparation for the PLATO Tutorial Logic (Compiler Version)." CSL Report I-130. Urbana: Coordinated Science Laboratory, University of Illinois, 1965.
4. Bitzer, D.; Lichtenberger, W.; and Braunfeld, P. G. "PLATO: An Automatic Teaching Device." IRE Transactions on Education E-4: No. 4, pp. 157-61; December 1961.
5. Bitzer, D.; Lyman, E. R.; and Suchman, R. "REPLAB: A Lesson in Scientific Inquiry Using the PLATO System." CSL Report R-260. Urbana: Coordinated Science Laboratory, University of Illinois, 1965.
6. Bitzer, Maryann. "Self-Directed Inquiry in Clinical Nursing Instruction by Means of the PLATO Computer-Controlled Simulated Laboratory." CSL Report R-184. Urbana: Coordinated Science Laboratory, University of Illinois, 1963.
7. Braunfeld, P. G. "Problems and Prospects of Teaching with a Computer." Journal of Educational Psychology, 55:201-11; 1964.
8. Braunfeld, P. G., and Fosdick, L. D. "The Use of an Automatic Computer System in Teaching." CSL Report R-160. Urbana: Coordinated Science Laboratory, University of Illinois, 1962.
9. Easley, J. A., Jr.; Gelder, H.; and Golden, W. "A PLATO Program for Instruction and Data Collection in Mathematical Problem Solving." CSL Report R-185. Urbana: Coordinated Science Laboratory, University of Illinois, 1964.
10. Lichtenberger, W.; Bitzer, D.; and Braunfeld, P. G. "PLATO II: A Multiple Student Computer-Controlled Teaching Machine." Programmed Learning and Computer-Based Instruction. (Edited by John E. Coulson.) New York: John Wiley and Sons, 1962. pp. 205-16.
11. Lyman, E. R. "A Descriptive List of PLATO Lesson Programs." CSL Report R-186. Urbana: Coordinated Science Laboratory, University of Illinois, 1965.

APPROACHES TO LIBRARY FILING BY COMPUTER

Jean M. Perreault

"The essence of library catalogue is arrangement of entries."¹

(a) There is no intention, in this paper, of providing anything like a new code of filing rules for use with the computer. Ted Hines and Jessica Harris have made a valiant and largely successful try at this task. It is recommended that you obtain it² and read it thoughtfully.

(b) Nor will this paper comment on the two classic American filing codes³ in such a way that the form subdivision for a subject-heading on this paper would read "—Commentaries," but rather in such a way as to give it "—Criticism, interpretation, etc." (c) Nor (as a final disclaimer) is this paper thematically concerned with the hope for code revision—though these pages come closer to such a treatment than to (a) or (b).

Instead, there will be an attempt (d) to present some of the intellectual or bibliographical problems involved in the notions of sorting and filing, and then (e) an outline of some of the tools and techniques which can be brought to bear upon their resolution. Together, these ideas should (f) make possible a rational basis for the evaluation of filing-code-revision suggestions.

The author is not entirely neutral in all this, as those who have read my paper "The Computer and Catalog Filing Rules"⁴ should know; but the following discourse will subject that paper to considerable retractatio. Primary influences in moving toward this revaluation were Lubetzky and Ranganathan. (To avoid confusion from the outset, the author does not take "library filing" to include the arrangement of documents, but only of their surrogates.)

Words—as anyone who has read, say, Finnegan's Wake, knows—can be a lot of fun. But they can also be quite a burden, particularly to whoever has the task of laying them end to end, or worse, one atop the other, in what could be called a vertical rather than a horizontal order. All discourse is the construction of "horizontal"

Jean M. Perreault is Lecturer, School of Library and Information Services, University of Maryland.

strings of words; the problem of filing is the construction of "vertical" stacks of words in discourse.

This last phrase should be emphasized because the filing problem, that of organizing stacks of discursions, is far more complex than that of simple sorting. The ideal, the situation which could enable fully adequate filing by ritualizations such as computers or filing clerks, would be the reduction of filing rules to sorting rules.

Filing and sorting could be distinguished thus: to sort is to set in order, to file is to interpolate into an already existing order of sorted items. But, common and easy to understand as this may be, it is far too superficial; even in the sorting process as defined here, there must necessarily arise interpolation—unless the items being "sorted" are somehow already in order. And, then, what does "in order" mean? We shall try to see later.

To sort is to put like with like. The term was used from the beginning of printing shop practice to mean the replacing of the elements of the broken up form, each element of which belongs with its own sort (type) in the case. "A" is a sort, "a" is another, and all the exemplars of each are put back into their appropriate sector(s) of the upper- or the lower-case, and this putting of sort with sort is sorting.

Words too can be sorted, just as can complex letter-forms. There is an interesting historical example of this cited in Douglas McMurtrie's history of printing The Book,⁵ where it is pointed out that the Chinese idea of moveable type (for printing in the Uigur language) was for the elementary pieces each to print a word, although Uigur is a language where words were built up from more elementary particles, namely letters. In our literal Western languages sorting is basically the ordering of the smallest elements, namely letters; but it is intriguing to try to imagine how an order of irreducible words could be set up. Let us not forget, however, that our own alphabetical (letter) order is every bit as arbitrary, although admittedly shorter than would be the (word) order of, say, the Mandarin vocabulary. Like the Chinese, a Western printer might well consider letter-complexes such as "á," "à," and "â" to constitute three additional sorts besides plain "a," but the breakdown into simple elements could go further, to "a," "´," "¨," and "ˆ", so that these accentual elements could be freely combined with other ordinary letters (as for instance "´" over "n" in Polish). By this technique, the total number of possible sorts is reduced, and this is only the logical extension of the reduction of the sorting of words themselves by means of the sorting of their elements, letters.

By "words", here, is meant each word—all the words of the dictionary taken individually. Discourse, where words are not taken individually, is quite another matter, because discourse cannot be standardized. It is this, in fact, that makes language fun, and it was

Joyce's refusal to accept the standardization of words, too, which makes *Finnegan's Wake* fun.

We can establish "sort-boxes" for everything from the elementary parts of an alphabet, or of all alphabets, up to the elementary parts (words) of a language. But how can the cross-over from this world of sorting to the quite different one of filing be accomplished?

First of all, since it has been postulated that it is non-standardizability that sets off the discursive from the sortably standardizable, we must try to see what it is, as far as order is concerned, that constitutes discourse. In answer to this I would suggest articulation, the same sort of phenomenon that lets a plurality of lines form various 2- or 3-dimensional figures by coming together at a variety of angles. Because of this variability, any two words can form as many patterns of discursive meaning as the language will allow—as is demonstrated, for instance, in the farce line "But you don't even know I'm alive" or in a great many puns.⁶

The usual way of handling these points of articulation between words is commonly expressed, among librarians, as "nothing before something," since the juncture between words is most commonly shown by a blank, a "nothing." This is true even in spoken language, where, although the blanks between words are not always clearly perceptible, at least those representing phrase-points normally are. These, indeed, are the historical antecedents in the development of our present-day rules of written-language punctuation.⁷ The reduction of the principle "nothing before something" to a purely sorting process is accomplished by treating the articulating blank as a part of the collating sequence, so that the sequence reads "blank, a, b, . . ."

But there are other kinds of articulation in "horizontal" discourse than that constituted by the blanks between words. There are, as mentioned, in spoken discourse: intonation, tempo, accent, etc.; in written discourse: punctuation, paragraphing, abnormal capitalization, italicization, etc. Although it is clearly possible to take these into account in our filing codes—by the device explained in "The Computer and Catalog Filing Rules"⁴ for instance—which device leads to phrase-by-phrase filing in addition to the word-by-word filing enabled by the use of the blank in the collating sequence—the real difficulty is the same as that encountered in mechanical translation and in classification for information retrieval with machines as the clerks. This central problem is ambiguity. It does not require much reflection to see that if each word in a language could denote one concept and one only, any word or words in that language could be easily enough translated into another similarly characterized language, provided that the conceptual underpinning and partitioning of the two languages was the same. That is, if there is no word in Eskimo for "to lie", then the English verb "to lie" cannot be adequately translated into Eskimo even if the structural similarity is there.

But in fact, languages are made up of words which are not univocal,⁸ and whose equivocacy becomes even more marked in discursive situations,⁹ since there the context sorts out the one sense among the several as appropriate. Take the phrases "a man of iron", "a feature of the land", "a man of Florence", "a book of John's." "Of" in these examples means successively "with the characteristic," "inherent in," "stemming from," and "belonging to." A precisely similar equivocacy can be seen in our marks of punctuation. (This is not, of course, to say that every preposition or mark of punctuation is equally burdened with equivocacy, but that some, such as "of" and the comma, are made to do a far greater share of the work than seems fair.)¹⁰

Now all this, in terms of the problem of filing, is incarnate in our desire to avoid purely mechanical rules and in our tendency to prefer an ordering of discursive elements in the same way that we articulate them in speech, in some way which transcends the purely temporal order of the sounds themselves. This tendency, or at least something faintly caricatural of it, is to be seen at work in the American Library Association and the Library of Congress filing rules. And lest we imagine that present-day Americans are alone in this tendency, bad or good, see the Prussian Instructions, part II.¹¹ In all these bodies of rules, we see a variety of solutions to the fact that in spoken discourse words can be variously emphasized, so that the usual order presumed as the basis of all filing is belied. This presumption is that the first element is always of higher importance than the second, and so forth.¹² (This point is given the name "Canon of Prepotence" by Ranganathan. Unfortunately, he forgets to indicate that, national as this position or unconscious policy may be, it is not the only one.)

This presumption should be examined, if only briefly. If there are two Dewey numbers, say 309.99 and 310, their filing order is determined by positional comparisons which accept the presumption mentioned above. The same is true as well for such a pair as 009 and 100, where the temptation would be strong for a naive person to put '1' earlier than '9' and not worry about the zeroes. But there is a different kind of number, one encountered even more frequently in real life: the integer as against the decimal number. In this case, two numbers are compared not on a first-to-first-element basis, and so forth, but in the exact reverse; here 30999 comes after, not before, 310.

Now although rules are needed for sorting of numbers, the verbal sorting/filing problem is far more difficult, or at least forms much the largest problem-sector in filing practice, except in the case of LC call numbers. There the presence of both decimal numbers and integers presents problems even to some librarians. And the interesting point is that sorting rules for words almost invariably operate according to the decimal sorting presumption. No one would

suggest that "mad" should come after "management" just because "-ent" is lower in the collating sequence than "mad" except, of course, in the organization of a rhyming dictionary. Yet there are cases where something close to the integer-sorting algorithm is at work with words, for instance where initial words such as articles are assumed to be of no value in the sorting order. (This tendency could, alternatively, be compared with the leading zeroes which get "9" in 009 to file ahead of "1" in 100. "The" at the head of a title is regarded as analogous to such ciphers.)

If sorting is the linear arrangement of simple or pseudo-simple inscriptions, and filing is the linear arrangement of semantic and discursive inscriptions, with all the difficulties implied, and if the goal is to make the problem of filing simple enough to be performed by clerks, human or mechanical, then we must seek for means to reduce these semantic-discursive aggregates to at least pseudo-simple ones. There are several tools or techniques for this task, ranging through quite a spectrum both in the complications of input required and in the output achieved.

If a clerk is given two entries to file relative to one another, one reading "Bible. English. 1964. Goodspeed." and the other reading "Bible. English. Knox.," the second files before the first not because letters, represented here by the "K" in "Knox" comes before the "1" in "1964", but because the Knox version, date unknown, would come among the first-order group of English (whole) Bibles with unspecified date-facets. (In such a case as this, a perhaps unexpected reflection of the rules for inverted file-order as derived from a citation-order can be seen quite clearly; see the Appendices for a fuller discussion.)

But for such an effect to be produced there must be instructions to the filing clerk that predict such possible variations, instructions such as "The facet after 'language' should be 'date'; if not, treat it as a blank date-facet, filed before all full date-facets." Essential in all this, if a clerk such as the computer is employed, is that each facet of such an entry be unambiguously labelled, so that the problem never occurs of the computer-clerk's not being able to "understand" the elements which it is called upon to manipulate.

Two instances of the same sort of discursive articulation can be seen in comparison (a) of such a pair of titles as The Man without Qualities and Man; Zephyr or Clod. Assuming that a dictionary catalog is being considered, the second of these will not be entered at all unless the explanatory sub-title is included, since otherwise it will be dropped because of a match with the subject heading apparatus. But if they are both entered and are to be filed relative to each other, the phrase-point introduced by the semi-colon cannot well be ignored lest the title be found only by someone who could imagine a title beginning (as if "man" were an adjective) Man Zephyr. . . . Thus, although

"w" in "without" precedes "Z" in "Zephyr", in the collating sequence, the phrase "Man", being shorter than "The Man without Qualities", should probably file first—just as a short word like "man" files before a long one like "management", and a short title like The Man before a long one like A Man and his Dog. It can also be seen in the order of (b) such a pair of entries as "Mass—Meditations" and "Mass (Canon law)." The parenthetical qualifier modifies the initial word in such a way as to make it represent a new concept, a longer one than the same word in its unqualified sense, even though the unqualified one is extended by a subdivision—and despite the priority of the "C" in "Canon" to the "M" in "Meditations." While something more than pure "alphabetic" or "symbolic" order is desired, there is an ineluctable dependence of all orders upon that which is "matter" to their "form"—in this case the ":", "—" or "()." "

Thus there is a definite justification for the traditional attitude of librarians that absolute adherence to the letters (or symbols) alone of an entry is not enough to generate an intelligible and useful file. "Order" is that which, hopefully, is generated by a file; just what state of affairs it is that can be characterized by "in order" is not clear.

Fugmann defines order as "The meaningful contiguity (Beieinandersein) of the parts of a whole."¹³ But this seems to imply perhaps too forcefully the naturalness of order; an organ in a body has such a part to play in the whole, but a word in a list of words is not related thus to the others. A bit more neutrally then, one might define order as "being where one ought to be." Thus "2" ought to be after "1", and "3" next; but the reason for "b" being after "a" is not quite the same. The order of letters is arbitrary, not as with numbers, natural. Yet even this last is not entirely true, since the number sequence could, if not radix-10, be quite other than is "naturally" expected. Thus it can be seen that it is finally the concepts referred to or represented by (1) the numbers not as mere digits, but as real numbers or (2) letters not as mere digits, but as elements of words, which are the basis for the order of a file—rather than that of a sorting, which depends entirely on "form" rather than "content." Or we could say, to follow the earlier terminology, "matter" rather than "form."

Even if there seems to be no real order except that of meaning, and if there remains a mixture of arbitrary and part-natural types, there is a common factor at work, namely that they both make prediction possible. The difference between sorting and filing comes sharply to the surface when the predictability to be found in each is sought. That of sorting is absolute, within the allowable limits of the decimal or integral presumption. But that of filing is only partial, (a) because of its dependence upon the meaning of the words (semantics), and (b) because of the meaning of the discourse constituted by these words as inter-related (syntax).

The essential ideas to keep in mind are:

(1a) Filing and sorting are not identical, nor do they refer to different stages of a single process. Sorting is putting like with like, whereas filing is the creation of a series of complexes by means of partial likenesses.

(1b) Discourse is what is to be filed, in libraries, and discourse is made possible by the articulation of semantemes (meaning-bearers) by syntax.

(1c) The result of this articulation, as well as of the character of many semantemes, is ambiguity. Even greater ambiguity characterizes a complex that is not articulated: thus our traditional and quite rational desire to retain as much as possible of the benefit of syntactic articulation.

(2a) There are two presumptions from which to choose in building a file: either the decimal or the integral. Numbers can be filed either way; words usually can be filed only decimally;

(2b) Means are needed to convert some words into "leading zeroes," as if the integral presumption were operative.

(3) The most general principles for filing are two: (a) "nothing before something," which can also be elaborated into the principle of inversion of the citation-order, and (b) order as predictability.

Some Tools and Techniques of Order Building

There can be two polar-antithetical positions taken with regard to the order of a file of complex verbalisms: the purely symbolic, and the purely eidetic. There are variations within each such pole, of course, so that within the adherents of eidetic order, there are (a) those who want nothing but classification, "Out with words altogether!", (b) those who want chain-indexing and the like, and (c) those who favor a strictly subjective or semantic interpretation of the words and punctuations, especially of subject headings. Also, within the adherents of symbolic order, there is the traditional disagreement over the recognition of the blank: (a) recognizing it yields word-by-word order, (b) ignoring it yields letter-by-letter order. Nor are these all the possible shades of divergence.

The scope of the present inquiry must accordingly be restricted, lest the luxuriance of detail make the presentation of the polarity unintelligible; the appendices, however, contain additional matter. Classification as an alternative to the subject heading catalog will not be considered even though I favor it myself. Nor will the chain-index. Nor will letter-by-letter filing be admitted as an acceptable device. There still remains a polarity between the symbolic and the

eidetic; and, as stated earlier, what should be sought is the reduction of filing rules (as called for by the eidetic nature of the items being manipulated) to sorting rules (as called for by the symbolic embodiment of these items).

The reason for this reduction is the need for univocity. The computer "knows" nothing; it cannot, as Fairthorne quips, find just the important words, because for a word to be important it must mean something. The computer can only manipulate words; meaning means nothing to it. So the equivocacy inherent in discourse must be avoided; and, if this can be accomplished, the computer is left with things to manipulate that it can handle (without needing to understand their meanings).

The two polar-antithetical positions then refer to the structure of the file that is to be generated. As to the means that are to lead to these positions, there are at least three, from which each position can choose at least two as its means of embodiment. The three means are (a) the collating sequence as such, (b) the numerical-interpolation device, and (c) the "programming" of the cataloger.

If there is a sample set of complexes to file, such as "WELLS (SOMERSET)" [a place as subject], "Wells (Somerset)" [a place as author], "Wells of despair" [a title], "WELLS, HERBERT GEORGE" [a person as subject], "Wells, Herbert George" [a person as author], "WELLS" [a topical subject], and "WELLS, OIL" [another topical subject], there can be two classes of order set up amongst them in accordance with the eidetic or the symbolic positions:

EIDETIC₁

Wells, Herbert George [peA]
WELLS, HERBERT GEORGE [peS]
Wells (Somerset) [plA]
WELLS (SOMERSET) [plS]
WELLS [tS]
WELLS, OIL [tS]
Wells of despair [T]

SYMBOLIC₁

WELLS [tS]
Wells, Herbert George [peA]
WELLS, HERBERT GEORGE [peS]
Wells of despair [T]
WELLS, OIL [tS]
Wells (Somerset) [plA]
WELLS (SOMERSET) [plS]

As mentioned, variations are possible within each of these polar positions; for instance, within the eidetic position all authors, subjects, and titles can be segregated (the divided catalog approach); or, within the symbolic, distinction between single- and double-blanks can be observed, giving:

EIDETIC₂

Wells, Herbert George [A,pe]
Wells (Somerset) [A,pl]
WELLS, HERBERT GEORGE
[S,pe]

SYMBOLIC₂

WELLS [first word followed by
many blanks]
Wells, Herbert George
WELLS, HERBERT GEORGE

WELLS (SOMERSET) [S,pl]

WELLS, OIL [first word
followed by double-
blank]

WELLS [S,t]

Wells (Somerset)

WELLS, OIL [S,t]

WELLS (SOMERSET)

Wells of despair [T]

Wells of despair [first word fol-
lowed by single-blank]

Now, consider the means for achieving these various results. The collating sequence, (a) may have too much expected of it to manage the eidetic orders, at least without nominal assistance from (c), the "programming" of the cataloger. The computer, in other words, can set first in order "Wells, Herbert George" as author and then as subject, because there is a difference in the characters of which each is composed—just as the filing clerk can tell apart black entries for authors and red entries for subject. But it can scarcely be expected to put places ahead of topics unless it has available a ridiculously large table of place names for reference. And that, in any case, is dependence on the cataloger too, as is depending on the tracing-group number-families to distinguish authors, titles, and subjects.

The numerical-interpolation device, (b) on the other hand, can achieve any desired order at all, but while it may be considered reasonable for authors and subjects—which can be counted on to return with some degree of frequency¹⁵—it is assuredly not so with titles, which can be counted on to be almost invariably varied, thus requiring a new number for almost every document. The numerical-interpolation device seems to me to have no place in the symbolic position, which by definition should not call upon any outside assistance, at least not beyond that which the cataloger does anyway, such as assignment of tracing-group number-families. (It does, however, have a very definite place in several clerical operations such as the periodical check-in file.)¹⁶

A tiny matrix can be constructed to show the intersection of these two sets of ideas (positions and means), with checks indicating mutual appropriateness:

		EIDETIC	
		/	SYMBOLIC
COLLATING SEQUENCE ALONE			X
NUMERICAL INTERPOLATION	X		
'PROGRAMMED' CATALOGERS	X	X	

The "programming" of the cataloger, (c) the only one of the three devices necessarily involved with both the eidetic and the symbolic positions, enters unrecognized even in the apparently pure collating-sequence technique as it is what distinguishes between otherwise identical entries. But it has a larger role to play in

situations of a higher complexity than those shown in the sample, such as the quasi-facet structure of Bible entries. (For greater detail, see the Appendices and the end of this section.)

It is well to detail the actual operations required by each of these three techniques before attempting to proceed to an evaluation of results and a recommendation of goals. The collating sequence might, if assumed to encompass a sufficiently wide gamut of discriminations, achieve even fairly good results in the eidetic as well as the symbolic ordering of entries. However, it has been my observation* that the letters of the alphabet can be called self-justificatory insofar as they either transcribe the given elements of the document or embody the controlled vocabulary decisions of the cataloger. This distinction between the a-scriptive and the de-scriptive is brought out more fully in my paper "Documentary Relevance and Structural Hierarchy."¹⁷ Our cataloging rules, contrasted with the policies of analytical bibliographers who treat punctuation and the like as being as sacrosanct as the words of the transcription, take an unfortunately cavalier attitude about punctuation. The rules avoid a clear-cut decision between two attitudes, one of which would lead to leaving the punctuation just as it is found (descriptive), the other to transforming given punctuation into functional (controlled, ascriptive). Thus, what the computer is fed, in terms of guidance for the application of the collating sequence (which, other than the alphabet and the numeral digits, should not be considered as frozen forever by the hardware manufacturers), is far from the kind of thing that can help in the decision, for instance, that parenthetical expressions like "(Somerset)" create a place sub-group. To do this, the parenthesis would have to be used for such a meaning and no other. As mentioned before, codes of descriptive cataloging do not bind their practitioners to a thoroughgoing usage of any particular punctuation symbol for any particular function. Nor, indeed, do codes for the rendering of headings. To insist on this consistency of usage would be, to some extent, further "programming" of the cataloger.

The two alternatives in the use of the collating sequence alone as a filing device—and note that such a device does indeed reduce filing to sorting—are (a) the superposition of the collating sequence onto present descriptive and entry practice, or (b) its superposition onto descriptive and entry practice thoroughly re-worked for the sake of such results. The operative presumption here is decimal, so much so that even integral numbers would have to be fitted out with leading zeroes (for instance, in dates of the Christian era earlier than A.D. 1000). The computer can then operate upon these atomized elements

*In the course of an investigation of the permissible compression of book-titles for the Bro-Dart Foundation. See Appendix IV.

(digits) of complex verbalisms, filing them by sorting them. But it must be noted too that unexpected difficulties can easily arise unless cataloging usage is modified, primarily in the direction of a far more stringent attitude about formats and the like.*

Use of the standard IBM collating sequence, for instance, in which the punctuations in our sample (comma, parentheses, blanks) are accepted as listed, gives (since the order built into the hardware is: blank, comma, opening parenthesis) SYMBOLIC₂. If policy were to dictate, instead, SYMBOLIC₁, it would be necessary to treat these various punctuations as blanks to be ignored in the creation of the "sort-tags." These then look (with "real" blanks indicated as "b") thus:

```
WELLSbbbbbb...
WellsbHerbertbGeorge
WELLSbHERBERTbGEORGE
Wellsbofbdespair
WELLSbOIL
WellsbSomerset
WELLSbSOMERSET
```

The numerical-interpolation device, (b) as mentioned is hardly applicable to symbolic-order filing, since it assumes decisions about order on the part of the cataloger. These decisions, once made, allow the computer to proceed to a kind of second-order symbolic-order sort. In the earlier example, for instance, the policy decision could be (instead) that EIDETIC₁ was the desirable order, and that a span of numbers, say 5000-6000, be allocated to all entries beginning "wells." The seven entries would thus generate an interval of $1000/7 = 142+$, which can be rounded out to 140. The seven entries are accordingly assigned numbers to "fossilize" their policy-determined order:

5070	Wells, Herbert George	5490	WELLS (SOMERSET)
5210	WELLS, HERBERT	5630	WELLS
	GEORGE		
5350	Wells (Somerset)	5770	WELLS, OIL
	5910 Wells of despair		

Later additions to the file are assigned such numbers as will intercalate them in the place appropriate to the policy (for instance, "WELLS FAMILY," $5210 + 70 = 5280$). These numbers, once assigned, are either (a) automatically added to each entry by the computer from a look-up table (a program arrangement can provide for

*Daniel Gore's character, Melissa Spindrift, has a point in her fanatical insistence on obedience to trivial rules—at least in terms of the needs of computers. (See Reference 18.)

signalling the cataloger whenever an unassigned entry is input), or (b) input by the cataloger on the basis of a reference table, thus substituting for the verbal entry-form entirely.

The device of "programming" the cataloger, (c), as mentioned, is implied in every code of cataloging; but the needed specific concern for the enabling of filing-manipulation is present only accidentally, if at all.

Substantial additions, dictated by the policy desired in terms of variations within the eidetic position, would be required if this device were to be successful in reducing filing to sorting. For instance, in the example of the two Bible entries, what is necessary to place "Goodspeed" before "1964" would be the indication of facets: "⁶Bible. [5,4] ³English. [2] ¹Goodspeed." as against "⁶Bible. [5,4] ³English. ²1964. [1]." (The blank facets "5" and "4" stand for section and book, respectively.) An even more difficult example is the filing of names such as those beginning with "John," which can be followed by sobriquets, place-designators, given names, titles of nobility or sovereignty, etc. Two lists of "John" entries are given in the Appendices. Appendix I is a list of the subject-heading usage and ordering of such entries, and Appendix II is a list of main entry usage and ordering as drawn from LC printed catalogs. Analysis of these lists shows that no principle is employed throughout which can be expected to provide the predictability necessary to locate any single entry or group of entries without scanning the whole list. Some seemingly outlandish placements may be due to nothing more than individual clerical errors rather than to the following of rules, but it should be remembered that the more complex the rules (and the less they depend on pure ritual) the more likely are cases of forgetfulness. However, in the filing of names of sovereigns, for any desired eidetic order to be achieved, it is necessary not only to indicate empty facets (as with the Bible entries), but to label each facet with a weighting factor not determined by its position in the entry as written, in accordance with such a citation order as "Name. Rank. Area of sovereignty. Number," as against the written order "Name. Number [and sobriquet]. Rank. Area of sovereignty". This is not to say that such relatively simple-minded means are capable of setting right every problem of filing order. Note, for instance, the resumption of a second sub-alphabet in the main entry list at "John Alcober." Several entries that follow this one, although treated in accordance with a logical rule similar to that mentioned in the earlier example of "Mass (Canon law)"—namely that a longer first facet follows a shorter, even where the shorter is followed by a subdivision-word of higher alphabetic position—do not seem legitimately to be considered as longer facets. They do not since the usual subdividing facet of the first sub-alphabet is a sobriquet or a place-designator, such as "John the Fearless, duke of Burgundy" or "John of Cappadocia".

The lengthening element of the supposedly longer first facet is similarly characterizable, for instance "John o'London, pseud." or "John Chrysostom, Saint" (Chrysostom is not a name, but a sobriquet meaning "golden-tongued.")

Implicit in these anomalies are the data that can lead to an important distinction between two aspects of the supreme desideratum, predictability. This distinction will be more fully discussed shortly.

An Attempt at Evaluation

Many solutions to these problems have been proposed, but if they do not recognize and control the actuality of the difficulties encountered in searching and in cataloging, they can be called "relative-ly simple-minded." Nor is it true that all such problems are necessarily real ones. A real problem in this area is one that not only presents us with an intricate tangle of circumstances, but one where-in a real need is prevented from being met. What are the real needs here? The most important, as can be seen from some of the foregoing, is predictability. To need to find a particular sovereign-person entry but to be unable to tell where to find it, even if the words used in it are known, is a need prevented from fulfillment. The same is true where it is necessary to find a particular subject heading made up of several phrases, even if one knows (or is told syndetically) the precise style even of the punctuation between the phrases; or where it is necessary to find a class of entries that form a group (among others) of divisions of some particular subject-heading or corporate entry. If it is impossible to predict where such entries or groups of entries are located in the file, there is a real problem.

Basically, the difficulty is with the lack of system in the cataloging policies in use among librarians. As a somewhat more protracted example, a discussion of the subject-heading "Art, Byzantine" follows. According to the LC authority-document Subject Headings,¹⁹ the heading is among the fourth group of divisions of the heading "Art." The first of these sub-alphabets is that made up of (a) form-divisions such as "Art—Congresses" and "Art—Philosophy," as well as references from unused headings such as "Art—Negroes see Negro art." The second sub-alphabet is made up of (b) place names. If such a name were in use at present, we might expect to find "Art—Byzantium" here. The implication is, however, that this group is to comprise documents on the artifacts in a place, rather than the artistry characteristic of a place. The third sub-alphabet is somewhat of a hash, including inverted-phrase adjectival (c) styles-cum-periods like "Art, Abstract" and "Art, Baroque," (d) styles-cum-places like "Art, Cluniac" and "Art, Oriental," along with (e) miscellaneous

orientations like "Art, Commercial see Commercial art," "Art, Immoral," "Art, Municipal," and "Art, Regional." (The application of facet-analysis to the elements of this sub-alphabet reveals the full extent to which lack of system can become enshrined in the hearts of the unwary.) The fourth sub-alphabet, to which belongs "Art, Byzantine," is made up of inverted-phrase adjectival (f) ethnic-cum-place names like "Art, African," "Art, Celtic," and "Art, Greek" (to which an xx-reference from "Classical antiquities" is prescribed, leaving the user somewhat puzzled if he has in hand a document on modern Greek art). Here is found the complement to the second sub-alphabet, (b): such headings comprise the artistry characteristic of Byzantium, rather than artifacts to be found in that city or empire. The fifth sub-alphabet is composed of inverted-phrase adjectival (g) periods like "Art, Ancient," "Art, Renaissance," and "Art, Modern—20th century" arranged chronologically in defiance of their verbal form. The sixth sub-alphabet comprises uninverted phrases beginning (h) "Art and . . .," "Art in . . .," and the like.

Now let us assume a person in need of documentary information about Byzantine art, a person to whom that turn of phrase appears natural. The syndetic apparatus of LC subject headings will tell such a person that a quasi-classification under "Art" is the proper place to look for such information, rather than under the first-thought-of expression. But it is not pointed out to him, except implicitly, where such an entry can be predicted to be among all those beginning "Art". What is explicitly (but still not thematically) pointed out is that the form of the heading-to-be-sought is "inverted-phrase adjectival", and that the adjective is formed from a proper name (for the sake of argument, I am willing to grant that no one is ignorant of the general meaning and formation of "Byzantine.")

Notice that this user cannot assign the heading "Art, Byzantine" to the fourth (ethnic-cum-place) sub-alphabet under "Art" because there is nothing in the syndetic link to tell him this. More fundamentally he cannot because even though he might be aware of the groups generally to be expected under a broad heading like "Art" and (c) even of the specific order of the ethnic-cum-place group among the others, there remains a fundamental difficulty. He may very well be thinking of "Art, Byzantine" as meaning something quite different from that implied by its official place—as a style-cum-period or style-cum-region heading (analogous to "Art, Gothic") in the third sub-alphabet, or as a period as such in the fifth (analogous to "Art, Renaissance"). This difficulty is similar to that noted earlier with regard to some of the irregularities of order under "John": if a person does not know the precise significance of part of a heading, an eidetic order may well be such as to prevent him from finding it unless he takes the time and trouble to look through all the headings beginning "John" or "Art". The difficulty arises because he does not

know the part of the heading's precise significance as used by the cataloger or filer, not that he does not know its precise form (which the syndetic apparatus reveals to him). The case of "Art" is analogous to that of "John" in the way in which it modifies the desideratum of predictability.

Notice also that in a case of sorting of meaningless and "atomic" signs (whether digit-particles like accents, digits, syllables—as in spoken Japanese—or even words—as in Uigur) the order is entirely arbitrary just because there is no meaning in their seriality. (The numerals "1" through "9" do not fall under this stricture except insofar as we presume the radix "10," and they are a paradigm then of both sorting and filing.) In an arbitrary sorting-order there is absolute predictability of the location of "q" between "p" and "r." The seriality of the whole alphabet (as governed by the Anglo-Saxon "arbitration") makes it possible to go from any entry-point in the series to the desired point: if one enters at "k", one knows that he must go "downstream" about half-way to the end of the series to hit "q."

In a file there is seriality too; but absolute predictability of the kind just mentioned requires an equally stringent application of arbitrary rules and an equally absolute awareness of the precise form of the sought sign-complex. It is the task of the syndetic apparatus to furnish the searcher with this absolute awareness because of the vagaries that seem inevitable in the cataloger's choice and rendering of entries. The seeker for "John Chrysostom, Saint" may well not be able to predict the location of such an entry precisely because he is aware of the fact that "Chrysostom" is a sobriquet, while the cataloger or filer is not.

A Conclusion of a Kind

It is easy enough to pick out flaws in any system, and far more so in any body of rules such as we are familiar with for filing, because none of them is really a system. I am not unsympathetic to the various attempts to provide the means for ordering a file of complex verbalisms. I am, however, convinced that only to those who look the difficulties full in the face and recognize the primordia of the concept of order in its various manifestations can there come hope of being able to devise or even to accept improvements.

The need for predictability is paramount, but there is not just one predictability to be sought for. In terms of the bifurcation "eidetic/symbolic" these insights appear: (a) symbolic order makes possible the predictability of entries known in precise detail, whereas (b) eidetic order makes possible the predictability of entries known in terms of their meaning. For instance, entries beginning

"M," "Mac," and "Mc," either in a separate alphabetic sector before or after all other "M. . ." entries or in a sub-sector of "M. . ." (as if all spelled "Mac"), make possible finding entries like "McClare," "MacClennan," and "M'Clare" in that order, without the searcher needing to advert to the precise form of the prefix. This is an example of the advantage of eidetic order. One advantage of the symbolic order is that the searcher for "Art, Byzantine" could afford to ignore the membership of this heading in one of the sub-alphabets. Instead, he could be assured that all comma-subdivisions of any heading like "Art" would be in a predictable place, distinct from double-dash-subdivisions, full-stop-subdivisions, parenthesis-subdivisions, or any other such.

Can we afford to sacrifice one such advantage to gain the other? I do not believe so; and I believe that a combination of two of the techniques outlined earlier, namely the collating sequence and the "programming" of the cataloger, can effect the solution desired without sacrifice of either advantage. These elements of a putatively adequate filing system are both indispensable anyway, since all filing implies (a) ritually sortable elements to constitute its filed complexes, and it is the collating sequence that makes such ritual sorting possible. It also implies (b) semantic content to these complexes, and the complementary control of them by the application of bibliographical control by the cataloger.²⁰ If these elements are combined with foresight, there can arise these two distinct and important results: (a) an eidetically determined filing order symbolically sortable, and (b) a flexible computer-contained surrogate capable of serving as a basis for a wide variety of output purposes,²¹ both ritually manipulable. By such a combination of devices, which would imply the tagging of a carefully determined totality of the sub-elements of the whole complex of descriptive elements, there could be brought about any order desired, as long as such a desired order was based on ritually consistent principles, and any degree of abbreviation of the entry as a whole. The main point to grasp here is that the greater the number of uniquely tagged sectors, the greater the flexibility for both these purposes. These sectors will not be easy to enumerate, nor will it be a simple matter to devise unique codes for them. It is probably possible only with the use of mutually modifying codes.

For instance, the combination of eidetic and symbolic order, through the conscious cooperation of the cataloger and the collating sequence, could give this order to the sub-alphabets of "Art": "Art—Congresses" and other forms, "Art: London" and other places in which. . . , "Art 'Gothic'" and other styles etc., "Art; Byzantine" and other places of which. . . , "Art/1901—" (=20th century modern) and other periods, "Art and . . . , Art in . . ." and other uninverted phrases, and finally, though not given (nor needed here) in Subject Headings, "Art (Qualified)." (In such an arrangement the punctuation used must

not be allowed to be used for any other purpose, at least not in subject-headings. This is what I mean by mutually modifying codes.) Thus the heading "Art, Byzantine" would either become "Art; Byzantine," with see-references from "Art 'Byzantine'" and "Art/[Byzantine period]" and (perhaps) "Art: Byzantium"—to force all usages of the inverted phrase to be located at the putatively most helpful place, or there could be entries allowed under each of the three, depending upon the emphasis of the document being cataloged (ethnic/geographic, stylistic, or temporal), with see-also references connecting them.

The steps to such a solution must include (a) comparative analysis of the various eidetic and symbolic orders advocated, in the light of basic principles such as have been enunciated here, (b) tabulation of the sectorings necessary to achieve all these orders—insofar as they can be seen to be truly desirable in terms of the real problems of searching—from a single input-record, and (c) checking to see whether additional sectoring is necessary to enable selective abridgements. This ambitious program does not, of course, mean that we cannot achieve improvements right now. Already, in the design of various computerized cataloging systems (as at Florida Atlantic University,²² the Ontario New Universities Library Project,²³ the University of California at Santa Cruz,²⁴ and the medical libraries' union catalog at Yale, Harvard, and Columbia Universities²⁵), various systems of use of computers in filing have been worked out. These may have been less than ideal solutions, but if they were, it is primarily because of the ad hoc nature of the attempts. They were not worked out as consciously partial solutions of the problem as a whole. The combination of a careful re-examination of our prevalent attitudes on preferred eidetic order(s), of the introduction, where necessary, of flexibility in the collating sequence, of the development of the codifications which can make cataloging rules bear along with them all that is needed for the accomplishment of the two central purposes—this is what ultimately must be sought, even when we wish to make only small practical improvements.

REFERENCES

1. Ranganathan, Shiyali Ramamrita. Heading and Canons; Comparative Study of Five Catalogue Codes. Madras, S. Vigwanathan and Sons, 1955, p. 60.

2. Hines, Theodore C. and Harris, Jessica L. Computer Filing of Index, Bibliographic and Catalog Entries. Newark, Bro-Dart Foundation, 1966.

3. A.L.A. Rules for Filing Catalog Cards. Prepared by a Special Committee, Sophie K. Hiss, Chairman. Chicago, American Library Association, 1942.

U. S. Library of Congress—Processing Department. Filing Rules for the Dictionary Catalogs of the Library of Congress. Rev. ed. of the 1945 Filing Manual. Washington, Library of Congress, 1956.

4. Perreault, Jean M. "The Computer and Catalog Filing Rules," Library Resources & Technical Services, 9:325-331, Summer 1965.

5. McMurtrie, Douglas Crawford. The Book; The Story of Printing and Bookmaking. 3d rev. ed. New York, Oxford University Press, 1943, pp. 95-97.

6. Cf. e.g. Robinson, Jane J. The Transformation of Sentences for Information Retrieval. Santa Monica, California, The RAND corporation (P-3243), December 1965 (Paper presented at the 1965 Congress of International Federation for Documentation (FID), Washington, D. C., October 1965).

7. See e.g. MacLuhan, Herbert Marshall. The Gutenberg Galaxy; The Making of Typographic Man. Toronto, University of Toronto Press, 1962, pp. 159-161.

"Applied knowledge in the Renaissance had to take the form of translation of the auditory into visual terms, of the plastic into retinal form;" note in particular his quotation from Father Ong (p. 160): "The spatial processing of sound by means of the alphabet is not enough. Printed or written words themselves must be deployed in spatial relationships, and the resulting schemata thought of as a key to their meanings."

8. Hjelmslev, Louis. "Prolegomena to a Theory of Language," Translated by Francis J. Whitfield In Indiana University Publications in Anthropology and Linguistics, Vols. 1-10 (1948-54): Memoir 7, pp. 29-38.

He points out here the non-cotermineity of the individual members of two groups (in English and Welsh) of words—and again, of three (in Danish, German, and French)—which, collectively, are coterminous.

9. Ceccato, Silvio, et al. Linguistic Analysis and Programming for Mechanical Translation. New York, Gordon and Breach, 1961, p. 59 ff.

10. Lubetzky, Seymour. Letter dated Oct. 13, 1964.

This is forcefully pointed out (in answer to my paper cited in reference 4), in a passage whose import I have attempted to integrate with my own prevailing conviction of the supremacy of the collating sequence (i.e. of sorting as the ideal ritual to make possible

computer-filing); he argues that "the basic difficulty lies in our use of the same symbols—the dash or the comma—for various types of subdivisions, and I believe that the development and use of a system of symbols to designate and distinguish the different types of subdivision would add a new dimension to the organization of the catalog and contribute materially to its intelligibility and effectiveness." This hope stems from his conviction that "to theorize that filing is implicitly alphabetization, not classification, is quite erroneous and misleading." I am not willing to go that far, at least terminologically, but the same insight is stressed here that motivated Ranganathan's statement quoted at the beginning of this paper.

11. The Prussian Instructions; Rules for the Alphabetical Catalogs of the Prussian Libraries. Translated from the Second Edition, Authorized August 10, 1908, with an Introduction and Notes by Andrew D. Osborn, Ann Arbor, U. of Michigan Press, 1938, pp. 27-106 et passim.

12. Ranganathan, op. cit., pp. 59-63.

13. Fugmann, Robert. "Ordnung—oberstes Gebot in der Dokumentation," Nachrichten für Dokumentation, 13:121, Number 3, Sept. 1962.

14. Needham, Christopher D. Organizing Knowledge in Libraries. London, Andre Deutsch, 1964, pp. 181-183.

15. According to an unpublished study made by Barbara Markuson of the Information Systems Specialist's Office of the Library of Congress, 47 percent of the authors examined had only one entry, 61 percent one or two, 66 percent one, two, or three; thus the hope for consistent return of author entries may be disappointed as well.

16. Cf. e.g. Srygley, Ted F. "Serials Record Instructions for a Computerized Serial System," Library Resources & Technical Services, 8:248-256, Summer 1964.

Schultheiss, Louis A. "Two Serial Control Card Files Developed at the University of Illinois, Chicago," Library Resources & Technical Services, 9:271-287, Summer 1965.

Stewart, Bruce Warren. A Computerized Serials Record System for the Texas A & M University Library. College Station, Texas A & M University Library, 1965.

17. Perreault, Jean M. "Documentary Relevance and Structural Hierarchy," Information Storage and Retrieval, 3:13-18, Number 1, 1966.

18. Gore, Daniel. "Encounter with a Cataloger," Library Resources & Technical Services, 9:363-366, Summer 1965.

19. U. S. Library of Congress. Subject Headings Used in the Dictionary Catalogs of the Library of Congress. 6th ed. Washington, D. C., 1957.

20. For discussion of the nature and function of bibliographical control in its various senses, see my paper: Perreault, Jean M. "On Bibliography and Automation or How to Reinvent the Catalog," Libri, 15:287-339, Number 4, 1965.

21. Cf. e.g. Buckland, Lawrence F. The Recording of Library of Congress Bibliographical Data in Machine Form. A Report Prepared for the Council on Library Resources, Inc. Rev. Washington, Council on Library Resources, 1965.

Note that Buckland's results by no means exhaust the possibilities, which can be made to extend from the full-standard analytical description to the briefest index entry.

22. See: Heiliger, Edward M., et al. "Florida Atlantic University," College & Research Libraries, 25:181-199, Number 3, 1964; and Perreault, Jean M. "Computerized Cataloging: The Computerized Catalog at Florida Atlantic University," Library Resources & Technical Services, 9:20-24, Number 1, 1965.

23. See: Bregzis, Ritvars. "Compatible Format for Machine Readable Record of Bibliographic Data," paper read at the 1965 Congress of the FID; Proceedings of the Congress to be published mid-1966; and also his "The Ontario New Universities Library Project-an Automated Bibliographic Data Control System," College & Research Libraries, 26:495-508, Number 6, 1965.

24. See: Computer Usage Company. Specification for an Automated Library System. Palo Alto, 1965.

25. See: Kilgour, Frederick G. "Development of Computerization of Card Catalogs in Medical and Scientific Libraries." In Herbert Goldhor, ed., Proceedings of the 1964 Clinic on Library Applications of Data Processing. Urbana, Illinois, University of Illinois Graduate School of Library Science, 1964, pp. 25-35.

ADDITIONAL REFERENCES

Graduate Research Center of the Southwest. Automation and Inter-Library Information Systems. A study sponsored by the Graduate Research Center of the Southwest, 1:24-105, Dallas, 1964.

Weinstein, Edward A., and George, Virginia. "Notes Toward a Code for Computer-Produced Printed Book Catalogs," Library Resources & Technical Services, 9:319-324, Number 3, 1965.

Appendix I

A list of subject heading uses of "John. . .," from: Library of Congress Catalog; (Books: Subjects) a Cumulative List of Works Represented by Library of Congress Printed Cards, 1955-1959, 11: 580-581. Patterson, New Jersey, Pageant Books, Inc., 1960.

JOHN BERCHMANS, SAINT	JOHN (BOOK OF THE NEW TESTAMENT)
JOHN BOSCO, SAINT	JOHN, EPISTLES OF (BOOK OF THE NEW TESTAMENT)
JOHN CAPISTRAN, SAINT	JOHN BROWN RAID, 1859
JOHN CLIMACUS	JOHN CRERAR LIBRARY, CHICAGO
JOHN FISHER, SAINT	JOHN DICKINSON AND COMPANY, LTD.
JOHN FREE	JOHN GOFFE'S MILL, BEDFORD, N. H.
JOHN, KING OF ENGLAND	JOHN HANCOCK BUILDING, SAN FRANCISCO
JOHN OF GAUNT, DUKE OF LANCASTER	JOHN LEWIS PARTNERSHIP, LTD.
JOHN OF JANDUN	JOHN MORE ASSOCIATION
JOHN OF KRONSTADT, FATHER	JOHN MUIR TRAIL, SIERRA NEVADA MOUNTAINS
JOHN OF LEYDEN	JOHN NEWBERY MEDAL BOOKS
JOHN OF ROQUETAILLADE	JOHN RYLANDS LIBRARY, MANCHESTER
JOHN OF SALISBURY, BP. OF CHARTRES	JOHN THE BAPTIST'S DAY
JOHN OF THE CROSS, SAINT	JOHN WOODMAN HIGGINS ARMORY, WORCESTER, MASS.
JOHN XXIII, POPE	
JOHN, PRESTER	
JOHN, SAINT, APOSTLE	
JOHN THE BAPTIST	
JOHN VIANNEY, SAINT	
JOHN, OTTO	

Appendix II

A list of author uses of "John. . .," from: A Catalog of Books Represented by Library of Congress Printed Cards, Issued to July 31, 1942, 76:616-621. Ann Arbor, Michigan, Edwards Brothers, Incorporated, 1944.

John III, <u>King of Portugal</u>	John <u>de Bordeaux</u>
John, <u>abbot of Ford</u>	John <u>de Britto</u>
John (bar manager)	John <u>de Halton</u>
John, <u>bp. of Lincoln</u>	John <u>de Pontissara, bp. of Winchester</u>
John, <u>bp. of Nikiu</u>	
John, <u>Brother</u>	John <u>de Sandale, bp. of Winchester</u>

John, <u>de Taxster</u>	John <u>Wallensis</u>
. . .	John <u>Alcober</u>
John <u>of Burgundy</u>	John <u>Ambrose, father</u>
John <u>of Cappadocia</u>	John <u>Amundesham</u>
John, <u>of Damascus, Saint</u>	John <u>Baptist, archduke of Austria</u>
John <u>of Gaddesden</u>	John <u>Boston, of Bury</u>
John, <u>of Garland</u>	John <u>Chrysostom, Saint</u>
John <u>of Gaza</u>	John <u>de Burgh</u>
John <u>of Genoa</u>	John <u>de Oxenedes</u>
John <u>of Halton, bp. of Carlisle</u>	John <u>Francis Regis, Saint</u>
John, <u>of Hildesheim</u>	John <u>Gabriel, sister</u>
John <u>of Hoveden</u>	John <u>Gualbert</u>
John <u>of Ire'land</u>	John <u>Henry, pseud.</u>
John <u>of Jandun</u>	John <u>Joseph Laurence Charles</u>
John <u>of Lancaster, pseud.</u>	<u>Louis, archduke of Austria</u>
John, <u>of Nepomuk, Saint</u>	John <u>Mary of Reifenberg, father</u>
John <u>of Pian</u>	John <u>o'London, pseud.</u>
John <u>de Carpine</u>	John <u>of God</u>
John <u>of Ravenna</u>	John <u>of Gaunt, duke of Lancaster</u>
John <u>of Salisbury, bp. of Chartres</u>	John <u>of Saint Thomas</u>
John, <u>of Tynemouth</u>	John <u>Paul, pseud.</u>
John <u>of Wales</u>	John <u>Sackland</u>
John <u>of Worcester</u>	John <u>III. Sobieski, King of</u>
John, <u>Prester</u>	<u>Poland</u>
John, <u>son of Sabanis</u>	John <u>Zapolya, King of Hungary</u>
John, <u>the Evangelist</u>	John, <u>Alfred</u>
John <u>the Faster</u>	
John, <u>the Fearless, duke of</u>	
<u>Burgundy</u>	
John, <u>the Painter</u>	

Appendix III

Inverted Citation Order

In order to insure that whole comes before part, general before specific, the principle of citation order in classification is subjected to inversion. Citation order results in predictability among the individual concepts ascribed to a single document; inversion of citation order results in predictability among the complexes (i.e., the documents themselves or their surrogates).

Two techniques are displayed below, either of which can assure proper inversion. There must, however, be a decision ahead of time

as to the citation order upon which the inversions are to be based. Or, if there seems no clear preference as to citation order, filing order can be decided first, and the citation order can be derived as its inversion. In practice the two lines of attack may be exercised simultaneously.

TECHNIQUE A. With postulated facets "place," "period," and "form," of a subject Y, we can construct a matrix showing citation order, and then deduce all the patterns of occurrence:

	"place"	"period"	"form"	
Y			x	(document a)
Y		x		(b)
Y		x	x	(c)
Y	x			(d)
Y	x		x	(e)
Y	x	x		(f)
Y	x	x	x	(g)

The occurrences are in terms of foci from each facet, of course. For instance, if Y is "architecture," document (a) could be a bibliography of architecture, (b) a study of architecture in the 20th century, (c) a bibliography of architecture in the 20th century, (d) a study of the architecture of France, (e) a bibliography of architecture in France, (f) a study of 20th century architecture in France, and (g) a bibliography of 20th century architecture in France.

TECHNIQUE B. Alternatively, the citation order can be expressed as an (ad hoc) weighting device. If the facet "place" is to come first in the citation order, it should be accorded the greatest weight, "period" the next greatest, "form" the least, and empty facets zero. Thus "place" = 3, "period" = 2, and "form" = 1, so that document (a) receives the weight 001, (b) 020, (c) 021, (d) 300, (e) 301, (f) 320, and (g) 321, putting them into the same order as before.

These techniques are appropriate not only to subject classification, but to the cataloging code as well. For instance, facet analysis is implicitly present in the construction of anonymous-classic headings such as:

Bible. English. Goodspeed.

Bible. English. 1964.

Bible. N.T. English. 1964.

Bible. N.T. Galatians. English. Goodspeed.

Bible. N.T. Galatians. English. 1964.

Bible. N.T. Galatians. 1964. Goodspeed.

"part"	"book"	"language"	"date"	"version"
		x		x
		x	x	
x		x	x	
x	x	x		x
x	x	x	x	
x	x	x	x	x

The basic principle operating here is "nothing before something": the unspecified before the specified, the whole before the part, or the general before the specific. The desired effect is predictability—here, of finding the desired document or surrogate (while citation order makes the analysis of the document and the creation of the surrogate predictable).

Appendix IV

Compression of Titles

A sample of titles was chosen from the Catalog of Books Represented by Library of Congress Printed Cards . . . 1942, according to the following rule: the first English title in the third column or the next following page on which an English title appears, of pp. 100, 200, 300, etc. of vol. 1, the same from pp. 101, 201, 301, etc. of vol. 3, progressing similarly through each odd-numbered volume up to vol. 23; whereafter it was re-cycled from p. 100 of vol. 25, 101 of vol. 30, etc. This resulted in a random sample of 235 titles. English was the only language accepted, not because there is no problem with foreign titles, but because the probability of coming to a clear understanding of basic principles is higher under such restrictions.

Two simultaneous tabulations were made:

(A) Of types of main entry encountered, which yielded 11 title main entries, 10 corporate main entries, and 214 personal main entries. The smallness of the proportion of entries not of the "normal" pattern, i.e., of a title distinctly subordinate to a personal responsibility, is significant, since the form of title in such "normal" cases is not as crucial in the finding operation as in corporate main entries or in title main entries. (It was not tabulated whether some of the "odd" 21 might, by a different decision, have been considered personal main entries, nor whether any of the "normal" 214 might have been considered "odd" by a different code for choice of entry.) It is perhaps interesting to note that none of the other main classes of main

entry occurred, namely anonymous classics or uniform headings.

(B) Of the structure of the titles themselves in terms of additional verbiage beyond the "core"-title, which yielded

- 1 alternative title as expansion of acronym,
- 1 sectional title,
- 2 mentions of contents,
- 4 initial words or phrases to be arbitrarily omitted as insignificant,
- 5 mentions of redaction,
- 7 initial possessives,
- 9 alternative titles introduced by "or,"
- 9 supplementary sub-titles,
- 11 mentions of occasion,
- 24 form sub-titles,
- 30 explanatory sub-titles,
- 39 mentions of scope, and
- 91 initial non-filing articles.

And, where a new class can be defined by the simultaneous membership of one title in two or more of the above in the same sector of the title, the tabulation yielded

- 1 explanatory/form/supplementary sub-title,
- 1 mention of occasion/explanatory sub-title,
- 2 alternative/explanatory sub-titles,
- 2 mentions of occasion/form sub-titles,
- 2 mentions of redaction/supplementary sub-titles,
- 4 mentions of scope/explanatory sub-titles, and
- 4 explanatory/form sub-titles.

The significant judgment to be made on these structural types is one of these three: (a) Is there a sector other than the "core"-title that should also be traced? (b) Is there a sector other than the "core"-title that should be omitted altogether at the time of compression? (c) Is there a sector other than the "core"-title that is necessary for the intelligibility of the "core"-title, but which should not be re-traced even when retained? Those classes that satisfy (a) are mentions of contents, sectional titles, and alternative sub-titles; those that satisfy (b) are explanatory and supplementary sub-titles, as well as some mentions of scope and some form sub-titles; those that satisfy (c) are mentions of redaction and the residual mentions of scope and form sub-titles. This, in terms of the sample, yields 12 sectors to be re-traced, ± 71.5 to be dropped, and ± 37.5 to be retained but not retraced. The remaining titles did not have such sectors, consisting of "core"-title alone.

Can the computer abbreviate titles according to these criteria? It might be that it could do so if standardized catalog-entry data were such as invariably to indicate these various functions—but the lack of

precision as to the various types of non-“core” sectors of titles, outside this investigation, makes it a priori impossible for such an indication to be expected. Something close to such an indication is implied by the practice of introducing alternative title in “or” by the invariable punctuation “; or,”. But examination of the titles in the sample leads to near-total lack of faith as far as any expectation of such consistency from LC cataloging is concerned. (It can, of course, be argued that this inconsistency is traceable back to the title-pages themselves, but any cataloger can easily dispute such an allegation, there being all sorts of arbitrary re-punctuations to be found in LC cataloging—although not for the sake of the establishment of any function such as automatic filing/tracing/abbreviation decisions.)

It is therefore recommended that four punctuation symbols be used in titles for purposes of functional sectoring, and that they not be allowed to be used in their common senses. They are:

(a) the colon (:) to indicate the re-tracing of mentions of content and of sectional titles,

(b) parentheses (()) to indicate the re-tracing of alternative subtitles,

(c) the semi-colon (;) to indicate that the subsequent data are not to be retraced or even retained at the time of compression, and

(d) the double-dash (--) to indicate the retention of the subsequent data, without re-tracing it.

One further symbol is to be used in a restricted way, the virgule (/); see below. This leaves enough punctuation symbols with their common usages (. , ? “ ‘ - * &) to avoid problems in the ordinary grammatical organization of the entry.

There is a large number of words which recur often enough to justify their invariable abbreviation, at least at the time of compression. The sample from which the following list is derived is certainly not large enough to lead us to think that they form an exhaustive tabulation, but they do form a sort of nucleus:

Amer(ican)	ess(ential)	p(ar)t
b(oo)k	extr(act) (acted)	pol(itical)
coll(ege)	gen(eral)	pr(inted) (inting)
collec(ted) (tion)	G(rea)t Br(itain)	publ(ished) (ication)
compl(ete)	h(an)db(oo)k	repr(inted) (inting)
conc(erning)	interpr(etation)	sect(ion)
cons(idered)	intro(duction) (duc-	sel(ected) (ection)
dev(elopment)	tory)	s(outh)
diss(ertation)	iss(ue)	tr(anslated) (anslator) (ans-
e(ast)	libr(ary)	lation)
ed(ited) (itor) (ition)	man(ual)	univ(ersity)
educ(ation)	mo(nth)	U(nited) S(tates)
Eng(lish)	n(orth)	w(est)
	org(anization)	y(ea)r

The standard abbreviations for the days of the week, the months of the year, the states of the Union should be used plus those for form-divisions given in Subject Headings.

Such words are not as frequent of repetition as are certain prepositional-phrase-introducers (which, while they do not each take up so much space, do so in the aggregate), such as:

as regards (a, an, the)
 at (a, an, the)
 by (a, an, the)
 for (a, an, the)
 from (a, an, the)
 in (a, an, the)
 of (a, an, the)
 to (a, an, the)
 with (a, an, the).

It is recommended that these phrase-introducers, together with the article following them, be programmed to be reduced to the virgule at the time of compression.

There are accordingly three steps in compression of titles, and they can be most satisfactorily accomplished in two stages in each of two possible ways:

(either)

(or)

- (a,b) functional punctuation inserted by the cataloger; abbreviation of frequent words by the cataloger
 (c) programmed excision of phrase-introducers (= '/')

- (a) functional punctuation inserted by the cataloger
 (b,c) programmed abbreviation of frequent words; programmed excision of phrase-introducers (= '/')

These techniques, when applied to the 235-title sample, result in compression of all titles to 40 or fewer characters except in 19 cases; the compressed titles thus truncated are somewhat less than intelligible in all but 8 cases. It seems unlikely that any further compression can be achieved except in terms of further truncation and of a direct increase in the number of unintelligibilities that result—except by the use of a longer list of frequent words to abbreviate.

The final problem is that of the elimination of words and phrases standing at the head of the title, rather than being sub-sectors coming after the “core”-title. There are three types of such initialisms to be eliminated: simple initial articles, initial possessives referring to the main entry, and insignificant phrases of introduction. Initial articles (in English) can be quite easily programmed out of the filing order in the same computer-pass which reduces phrase-introducers to virgules; however, the total elimination of such

articles in the compressed print-out is not recommended except where it is noticed (by program) that such retention will result in truncation of sectors neither re-traced nor to be eliminated, namely if the title is sectoried by a double-dash or is not sectoried at all. Of the 19 titles truncated at character 41, 12 began with articles, and the application of the ritual questions here recommended would have mollified the resultant unintelligibility by dropping the initial articles in 11 of those cases.

Initial possessives should always be eliminated from the very first if they refer to the main entry; there is only one case in the sample that does not reproduce the main entry, and must accordingly not be dropped. This "exception" occurs not so seldom in the real world, especially in humanistic and historical studies (e.g., "Aristotle's Contribution to . . .," by Barker, Mansion, or the like).

The 4 arbitrarily excised initialisms (type c) are characteristically equivalent to form sub-titles, but their position causes trouble in that the same symbol cannot be used to exclude them that was used for terminal exclusion of form sub-titles, namely the semi-colon. Accordingly, it is recommended that such initialisms, when recognized by the cataloger as accompanying a title that will result in being truncated by compression, should be excised from the very first, so as not to appear in the full title any more than in the compressed. Perhaps, as a concession, the brief lacuna-symbol (..) could be used to indicate such excision.

The titles that make up the sample from which these conclusions are drawn, first in raw form (with indication of structural type of sectoring) and then as compressed (with truncated parts underlined), are

History of King Richard the Second of England
HIST./KING RICHARD THE 2D/ENG.

Elementary mechanical drawing, for school and shop (explanatory/
scope)

ELEMENTARY MECHANICAL DRAWING--SCHOOL + SHOP

The Joseph Leidy commemorative meeting, held in Philadelphia,
Dec. 6, 1923 (occas.)

THE JOSEPH LEIDY COMMEMORATIVE MEETING; philadelphia,
Dec. 6, 1923

A relation of the great river of Amazons in South-America (scope)
A RELATION/GREAT RIVER/AMAZONS; s. amer.

Provincial society, 1690-1763 (scope)

Address to the people of Pennsylvania (title main entry)
ADDRESS/PEOPLE/PA.

War against war; or, The joys of peace (alternative)

WAR AGAINST WAR (the joys of peace)

THE JOYS OF PEACE

A manual of instruction in Latin on the basis of a Latin method. . .
(explanatory)

A MAN./INSTRUCTION/LATIN;/basis/latin method. . .

The Sumner assault. Speech of. . . (occasion, form)

THE SUMNER ASSAULT; speech of. . .

Hungary before, during and after the great war (scope)

HUNGARY--BEFORE, DURING + AFTER THE GREAT WAR

The Amaranth; a gift for all seasons (title main entry, supplementary)

THE AMARANTH; a gift for all seasons

Bulletin (corporate main entry)

The perfect tribute

Annable's grain calculator, showing at a glance the number. . . (init.,
explan.)

GRAIN CALCULATOR; showing/glance the number. . .

Queen Elizabeth

The chace

Arbor day planting in eastern states (title main entry, scope)

ARBOR DAY PLANTING--EASTERN STATES

The Arena (title main entry)

Report of audit (corporate main entry)

Farm experiments with fertilizers

FARM EXPERIMENTS/FERTILIZERS

Guide to the material on Swedish history in the Augustana college
library (scope)

GUIDE/MATERIAL/SWEDISH HIST.--AUGUSTANA COLL. LIBR.

Adelaide

Consular and other conventions between Austria-Hungary and Servia
(initialism)

CONVENTIONS/AUSTRIA-HUNGARY + SERVIA

Translation of the Mining law of Mexico, effective August 1, 1926. . .
(init., occas.)

MINING LAW/MEXICO; effective. . .

Progress in the science of education in the last twenty-five years
(scope)

PROGRESS/SCIENCE/EDUC.--LAST 25 YRS.

A collection of rare and valuable books

A COLLECTION/RARE + VALUABLE BKS. ('collection' not abbrev.,
because first word)

Q 39

The Barbers' journal (title main entry)

An introduction to the Italian language

AN INTRO./ITALIAN LANGUAGE

Marti, a story of the Cuban war (explanatory/form)

MARTI--A STORY/CUBAN WAR

A brief sketch of the history of the Catholic church in the island. . .
(init., scope)

THE CATHOLIC CHURCH IN THE ISLAND. . .

A sketch of the doctrine relative to commitments in bankruptcy

THE DOCTRINE RELATIVE/COMMITMENTS IN BANKRUPTCY

Sex, marriage and motherhood

SEX, MARRIAGE + MOTHERHOOD

Reconnaissance of the gold fields of southern Alaska with some
notes. . . (suppl.)

RECONNAISSANCE/GOLD FIELDS/S. ALASKA;/some notes. . .

The eight crooked trenches

All in it

History of the Forty-eighth Ohio vet. vol. inf. giving a complete. . .
(explan.)

HIST./48TH OHIO VET. VOL. INF.; giving a compl. . . .

Labor disputes and the President of the United States

LABOR DISPUTES + THE PRESIDENT/U. S.

The fundamentals of naval tactics

THE FUNDAMENTALS/NAVAL TACTICS

Joseph King Fenno Mansfield

Personal and literary memorials

Germany before the war

GERMANY--BEFORE THE WAR

Solar hydrodynamics

Commentaries on the law of England. . . the 19th ed. With the last
corrections of the author, and copious notes. By J. E. Hoveden
(redaction)

COMMENTARIES/LAW/ENGLAND--/NOTES/HOVENDEN

Co-operative movement in Russia (scope)

CO-OPERATIVE MOVEMENT--RUSSIA

The roads and road material of Indiana (scope)

THE ROADS + ROAD MATERIAL/INDIANA

Agricultural climatology of the United States compared with that of other parts of the globe (scope embedded in 'core')

AGRICULTURAL CLIMATOLOGY/U.S. COMPARED/OTHER PARTS/
GLOBE

Effects of various dietary deficiencies upon the morphology of the suprarenal gland

EFFECTS/VARIOUS DIETARY DEFICIENCIES/MORPHOLOGY/
SUPRARENAL GLAND (or, to force re-tracing of the second sector:)

EFFECTS/VARIOUS DIETARY DEFICIENCIES:MORPHOLOGY/
SUPRARENAL GLAND

MORPHOLOGY/SUPRARENAL GLAND

Psychology; a factual textbook (form)

PSYCHOLOGY; a factual textbook

The history of educational legislation in Ohio from 1851 to 1925 (scope)

THE HIST./EDUC. LEGISLATION--OHIO, 1851-1925

Boston investigator, devoted to the development and promotion of universal mental liberty (explanatory, title main entry)

BOSTON INVESTIGATOR; devoted. . .

The art of dressmaking at home and in the workroom (scope)

THE ART/DRESSMAKING--/HOME +/WORKROOM

Canada under British rule, 1760-1905 (scope)

CANADA UNDER BRITISH RULE--1760-1905

A history of Philadelphia, with notice of villages, in the vicinity (supplementary)

A HIST./PHILADELPHIA; / notices/villages,/vicinity

The drama and music in New South Wales (scope)

THE DRAMA + MUSIC--NEW S. WALES

The layman's handbook to Daniel

The LAYMAN'S HDBK./DANIEL

The Jacobins; an essay in the new history (form)

THE JACOBINS; an essay. . .

New reading room and libraries

NEW READING ROOM + LIBRS.

British standard specification for the dimensions of grinding wheels
and methods of attachment (initialism)

SPECIFICATION/DIMENSIONS/GRINDING WHEELS + METHODS OF ATTACHMENT

Lectures illustrative of various subjects in pathology and surgery
LECTURES ILLUS./VARIOUS SUBJECTS/PATHOLOGY + SURGERY

The Senate finance bill

Sanctuary, Sunshine house sonnets (form)
SANCTUARY--SUNSHINE HOUSE SONNETS

Lands of the Andes and the desert
LANDS/ANDES + THE DESERT

Catalogue of the library and prints of. . . of. . .
CATALOGUE/LIBR. + PRINTS/. . .

The crimson cutlass

The Pan-American conferences and their significance
THE PAN-AMER. CONFERENCES + THEIR SIGNIFICANCE

Summer tours (corporate main entry)

The blessed bees

Moses Greeley Parker, M.D.

A tentative bibliography of the belles-lettres of the Argentine republic
(scope)

A TENTATIVE BIBL./BELLES-LETTRES/ARGENTINE REPUBLIC

The third part of the Institutes of the laws of England: concerning
high treason, and other pleas of the crown, and criminal causes
(sectional)

THE 3RD PT./INSTITUTES/LAWS/ENG.: CONC. HIGH TREASON, +
OTHER PLEAS/CROWN
CONCERNING HIGH TREASON, + OTHER PLEAS/CROWN

English teaching in the Southwest; organization and materials for in-
structing Spanish-speaking children (scope/explanatory)
ENG. TEACHING--S.W.; org. + materials/instructing spanish-
speaking children

Christian salvation, a modern interpretation (explanatory)
CHRISTIAN SALVATION; a modern interpr.

Short texts from Coptic ostraca and papyri
SHORT TEXTS/COPTIC OSTRACA + PAPYRI

Meat purchasing a science

Suggestions for teaching primary reading and phonics
SUGGESTIONS/TEACHING PRIMARY READING + PHONICS

Saint Brigid of Ireland

North American ichneumon-flies, new and described, with taxonomic
. . . (suppl.)

N. AMER. ICHNEUMON-FLIES--NEW + DESCRIBED; /taxonomic +
. . .

The Greek herbal of Dioscorides; illustrated by a Byzantine, A.D. 512;
Englished by John Goodyer, A.D. 1655; edited and first printed,
A.D. 1933, by Robert T. Gunther (form, redaction, internal ref. to
main entry)

THE GREEK HERBAL; illus./byzantine, a.d. 512; eng./goodyer, a.d.
1655--ED. + . . .--ED. + 1ST PR., A.D. 1933/GUNTHER

Annelids from the Danmark expedition (scope)

ANNELIDS--DANMARK EXPEDITION.

Old Kensington palace and other papers

OLD KENSINGTON PALACE + OTHER PAPERS

Doderlein's Hand-book of Latin synonyms (initialism)

HAND-BOOK/LATIN SYNONYMS (first word not abbreviated)

Statement of the minister of the Dominican Republic, . . ., concerning
the incident of October last, in one section of the Haitian-
Dominican border (scope)

STATEMENT CONC. THE INCIDENT/OCT. LAST; in 1 sect./haitian-
dominican border

Historical sketch of the Congregational Church in Belchertown, Mass.,
from its reorganization, 114 years, with. . . (scope)

HISTORICAL SKETCH/CONGREGATIONAL CHURCH/BELCHER-
TOWN, MASS.; from. . .

Latin grammar

Madame Therese; or, The volunteers of '92 (alternative)

MADAME THERESE (THE VOLUNTEERS/'92)

THE VOLUNTEERS/'92

Jane Clegg; a play in three acts (form)

JANE CLEGG--A PLAY; /3 acts

Nicholas II, prisoner of the purple (explanatory)

NICHOLAS 2D--PRISONER/PURPLE

"The North Pole", a tale of the Arctic sea; a 100% tone and speaking
picture with songs, choruses and dances (explanatory/form/sup-
plementary)

THE NORTH POLE; a tale. . .

The margin of music

Two Tudor books of arms

TWO TUDOR BKS./ARMS (Numeral word not transliterated because initial)

An elementary geography; being also a key to the new series of outline maps (suppl.)

AN ELEMENTARY GEOGRAPHY; being. . .

Text-book of bacteriology

TEXT-BK./BACTERIOLOGY

Inventory of Memorials, presented, by the deputies of the Council of Trade in France, to the Royal Council, in 1701 (initialism, occasion, repet. of main ent.)

MEMORIALS, PRESENTED/ROYAL COUNCIL; /1701

Paris following the terror (scope)

PARIS--FOLLOWING THE TERROR

Romances of the law

ROMANCES/LAW

Map of southern Idaho and the adjacent regions (supplementary)

MAP/S. IDAHO; + the adjacent regions

Simplified spelling from the scholarly standpoint and with reference to the masses: the argument for simplified spelling (explanatory, alternative)

SIMPLIFIED SPELLING; /scholarly standpoint + /masses (THE ARGUMENT/SIMPLIFIED. . .

THE ARGUMENT/SIMPLIFIED SPELLING

A manual of procedure for the clinical laboratory

A MANUAL/PROCEDURE/CLINICAL LABORATORY

How I became a preacher

The cementation of iron and steel

THE CEMENTATION/IRON + STEEL

Methods of recording retail prices and measuring the cost of living in Italy (scope)

METHODS/RECORDING RETAIL PRICES + MEASURING THE COST/LIVING/ITALY

Principles of American forestry

PRINCIPLES/AMER. FORESTRY

Charter rights of the Greenville. . . railroad company (ref. to main entry)

CHARTER RIGHTS

Man-midwifery exposed and corrected; or, The employment. . .

(alternative)

MAN-MIDWIFERY EXPOSED + CORRECTED (THE EMPLOYMENT...
THE EMPLOYMENT. . .

Indians of yesterday

INDIANS/YESTERDAY

The ancestress, tragedy in five acts (form)

THE ANCESTRESS; tragedy. . .

Children's toys of bygone days; a history of playthings of all peoples
from prehistoric times to the XIXth century (explanatory/scope)

CHILDREN'S TOYS/BYGONE DAYS; a hist./ . . .

May dust

The story of mankind; science of anthropology--man's evolution--his
physical, mental and moral development (explanatory, alternative)

THE STORY/MANKIND (SCIENCE/ANTHROPOLOGY); man's . . .

SCIENCE/ANTHROPOLOGY

The book of electrical wonders

THE BOOK/ELECTRICAL WONDERS

The little schoolmistress`

The first (-second) book of word and sentence work; or, Easy steps
in spelling (inserted content, alternative)

THE 1ST-2ND BK./WORD + SENTENCE WORK (EASY STEPS/SPELL-
ING)

EASY STEPS/SPELLING

Map and aerial photo reading simplified

MAP + AERIAL PHOTO READING SIMPLIFIED

How not to treat Illinois soils

Case studies in the development of social attitudes

CASE STUDIES/DEV./SOCIAL ATTITUDES

The absorption of resonance neutrons by boron, chlorine, cobalt and
manganese

THE ABSORPTION/RESONANCE NEUTRONS/BORON, CHLORINE,
COBALT + MANGANESE

Our debt to the red man; the French-Indians in the development of the
United States (explanatory/alternative)

OUR DEBT/RED MAN (THE FRENCH-INDIANS/DEV./U. S.)

THE FRENCH-INDIANS/DEV./U. S.

A manual of the timbers of the world; their characteristics and uses
(explan.)

A MANUAL/TIMBERS/WORLD; their. . .

Assessment and taxation in Kansas (scope)

ASSESSMENT + TAXATION--KANSAS

The Italian navy in the world war 1915-1918 (scope)

THE ITALIAN NAVY--WORLD WAR 1915-1918

Body form in growing chickens

BODY FORM/GROWING CHICKENS

The present status of the certification of teachers in the United States (scope)

THE PRESENT STATUS/CERTIFICATION/TEACHERS/U. S.

The temperaments; or, The varieties of physical constitution in man, considered in their. . . (alternative/explanatory)

THE TEMPERAMENTS (THE VARIETIES/PHYSICAL CONSTITUTION/MAN; cons./mental. . .)

THE VARIETIES/PHYSICAL CONSTITUTION/MAN; cons./mental. . .

Report of the Central board. . . (corporate main entry)

Shakespeare's heroines; characteristics of women, moral, poetical, and historical (explan.)

SHAKESPEARE'S HEROINES; characteristics. . .

The memoirs of the Marquise de Keroubec (1785-1858) being extracts. . . (explanatory/form, reference to main entry)

THE MEMOIRS; being. . .

The Peacock farm

Technology, employment, and output per man in petroleum and natural-gas production (scope)

TECHNOLOGY, EMPLOYMENT, + OUTPUT/MAN--PETROLEUM + NATURAL-GAS PRODUCTION

The way of the West

THE WAY/W.

Poems; including The saint's tragedy, Andromeda, songs, ballads, etc. (contents)

POEMS; including: THE SAINT'S TRAGEDY: ANDROMEDA; songs

. . .
THE SAINT'S TRAGEDY
ANDROMEDA

A digest of the statutes of Arkansas, embracing. . . (explanatory/scope)

A DIGEST/STATUTES/ARK.; embracing. . .

The adventurous life of Count Lavalette, Bonaparte's aide-de-camp and Napoleon's postmaster-general (explanatory)

THE ADVENTUROUS LIFE/COUNT LAVALETTE--BONAPARTE'S. . .

Guy Livingstone; or, "Through" (alternative)
 GUY LIVINGSTONE (THROUGH)
 THROUGH

The book that gave to Iowa its name, a reprint (formal)
 THE BOOK THAT GAVE/IOWA ITS NAME; a repr.

Nutrition

Driven to bay
 DRIVEN/BAY

The perception of light and color
 THE PERCEPTION/LIGHT + COLOR

Morphological studies of the head and mouth-parts of the mature
 codling-moth larva, *Carpocapsa pomonella* (Linn.) (explanatory)
 MORPHOLOGICAL STUDIES/HEAD + MOUTH-PTS./MATURE
CODLING-MOTH LARVA; *Carpocapsa*. . .

Madame L'Orient's art of beauty (initialism)
 ART/BEAUTY

The American vine dresser's guide
 THE AMER. VINE DRESSER'S GUIDE

A manual of modern scholastic philosophy
 A MANUAL/MODERN SCHOLASTIC PHILOSOPHY

An oration, pronounced July 4, 1799, at the request of the inhabitants
 of the town of Boston, in commemoration of the anniversary of
 American independence (occasion)
 AN ORATION; pronounced. . . Boston--IN COMMEMORATION. . .
--IN COMMEMORATION/ANNIVERSARY/AMER.
INDEPENDENCE

The book of prescriptions, with notes on the pharmacology. . . (sup-
 plementary)
 THE BOOK/PRESCRIPTIONS; /notes. . . .

Adventures in interviewing
 ADVENTURE/INTERVIEWING

Catalogue of Marietta library (corporate main entry referred to)
 CATALOGUE

The horn-fly (*Haematobia serrata* Rob.-Desv.) (explanatory)
 THE HORN-FLY; *Haematobia*. . .

The law of insolvency in South Africa (scope)
 THE LAW/INSOLVENCY--S. AFRICA

The vegetation of twin island
 THE VEGETATION/TWIN ISLAND

Martin's natural history (initialism)
NATURAL HIST.

Upper night

Churches in lower Nubia (scope)
CHURCHES--LOWER NUBIA

The new State department

What is wrong?

Oil recovery investigations of the Petroleum experiment station of
the U. S. Bureau of mines

OIL RECOVERY INVESTIGATIONS/PETROLEUM EXPERIMENT
STATION/U. S. BUREAU/MINES

The olive branch; or, The evil and the remedy (alternative)
THE OLIVE BRANCH (THE EVIL + THE REMEDY)
THE EVIL + THE REMEDY

The doctrine of fascism
THE DOCTRINE/FASCISM

The law of suffrage and elections
THE LAW/SUFFRAGE + ELECTIONS

The archaeological collections (corporate main entry)
THE ARCHAEOLOGICAL COLLECS.

Rules and practice of the Supreme Court of South Africa, Transvaal
provincial division, Witwatersrand local division, and Appellate
division (contents)

RULES + PRACTICE/SUPREME COURT--S. AFRICA: TRANSVAAL
PROVINCIAL DIVISION: WITWATERSRAND LOCAL DIVISION:
APPELLATE DIVISION

TRANSVAAL PROVINCIAL DIVISION (a more desirable result would
WITWATERSRAND LOCAL DIVISION be to get each of these re-
APPELLATE DIVISION tracings as a continuation of
the opening phrase)

Suggested forms for internal financial reports of colleges and uni-
versities (scope)

SUGGESTED FORMS/INTERNAL FINANCIAL REPORTS/COLLS. +
UNIVS.

The war and South American trade, issued. . . by. . . (occasion, refer-
ence to m. e.)

THE WAR + S. AMER. TRADE; iss. . . .

Report of the secretary of the. . . on the census. . . 1861 (ref. to corp.
main ent.)

REPORT/CENSUS

Relation of. . ., tr. from the Spanish by Buckingham Smith (reference to main entry, redaction)

RELATION--TR./SPANISH/SMITH

Oaths and obligations of free masonry, as published to the world, by a convention of seceding masons, held. . . (title main entry, explanatory/occasion)

OATHS + OBLIGATIONS/FREE MASONRY; as publ./world. . .

A Gael's guide to Waterford and the Deise country

A GAEL'S GUIDE/WATERFORD + THE DEISE COUNTRY

The church of Rome: a view of the peculiar doctrine, religious worship, ecclesiastical polity, and ceremonial observances, of the Roman Catholic Church (expl.)

THE CHURCH/ROME; a view/(THE PECULIAR DOCTRINE, RELIGIOUS WORSHIP, ECCLESIAST. . .)

THE PECULIAR DOCTRINE, RELIGIOUS WORSHIP, ECCLESIASTICAL. . .

Poems. Dedicated to Thomas Moore, esq. (form)

POEMS--DEDICATED/THOMAS MOORE, ESQ.

The man with the lantern

THE MAN/LANTERN

Love in its extasie; or, The large prerogative (alternative)

LOVE/EXTASIE (THE LARGE PREROGATIVE)

THE LARGE PREROGATIVE

Export opportunities in South and Central America (scope)

EXPORT OPPORTUNITIES/S. + CENTRAL AMER.

An answer to the remarks of the Plymouth company, or (as they call themselves) the Proprietors of the Kennebeck purchase. . . on the plan and extracts of deeds

AN ANSWER/REMARKS/PLYMOUTH COMPANY; or. . . (LC has a tracing "Plymouth Company, 1749-1816. Remarks on the plan and extracts of deeds" which probably cannot be automatically extracted)

A handbook to the palace of Minos at Knossos, with its dependencies (scope, suppl.)

A HANDBOOK/PALACE/MINOS/KNOSSOS; /its. . .

Nutrition and the school lunch

NUTRITION + THE SCHOOL LUNCH

Some decisions on the Property acts

SOME DECISIONS/PROPERTY ACTS

Taxation in New York, 1924; the complete New York tax law. . .
(scope, explanatory)

TAXATION--N. Y., 1924; the. . .

Essays on the political history of the fifteenth, sixteenth, and seventeenth centuries

ESSAYS/POL. HIST./15TH, 16TH, + 17TH CENTURIES

Church music in the light of the motu proprio; a guide for the Catholic choir-master and organist (explanatory)

CHURCH MUSIC/LIGHT/MOTU PROPRIO; a guide. . .

D. Francisco Manuel de Mello

The dollar crop, and its effect upon labor, agriculture, commerce and manufacture

THE DOLLAR CROP, + ITS EFFECT/LABOR, AGRICULTURE, . . .

Reynold's code. Polyglot nautical telegraph for the use. . . (initialism, explan.)

CODE (POLYGLOT NAUTICAL TELEGRAPH; /use. . .

POLYGLOT NAUTICAL TELEGRAPH; /use. . .

Acts relating to the public schools of Rhode Island, with remarks + forms. (scope, supplementary)

ACTS/PUBLIC SCHOOLS/R. I.; with. . .

Rustic architecture

Fair exchange; a novel in the first person (form)

FAIR EXCHANGE; a novel. . .

Memoirs of Duke de Richelieu (reference to main entry)

MEMOIRS

Racing bits; a story of the oil-fields of Texas (form/explanatory)

RACING BITS--A STORY/OIL-FIELDS/TEX.

Floats for the so-called fattening of oysters

FLOATS/SO-CALLED FATTENING/OYSTERS

An address to the flocks of the reverend approvers of Blanco White's internal evidences against Catholicism

AN ADDRESS/FLOCKS/REVEREND APPROVERS/BLANCO WHITE'S:
INTERNAL EVIDENCES AGAINST CATHOLICISM

INTERNAL EVIDENCES AGAINST CATHOLICISM

Obstetrical notes for juniors (occasion)

OBSTETRICAL NOTES/JUNIORS

Canary breeding for beginners; a practical up-to-date guide (occasion/form)

CANARY BREEDING/BEGINNERS; A PRACTICAL UP-TO-DATE
GUIDE

St. Philip's church. Memorial poem (title main entry, form)
 ST. PHILIP'S CHURCH. MEMORIAL POEM (retention of original
 punctuation leaves whole as single field)

Annual report of the town (corporate main entry)
 ANNUAL REPORT/TOWN

The wayfarer; a novel (form)
 THE WAYFARER; A NOVEL

A bibliographical manual for the student of criminology (occasion)
 A BIBLIOGRAPHICAL MAN.--/STUDENT/CRIMINOLOGY (first word
 not abbrev.)

Syllabus of Davis' International law
 SYLLABUS/DAVIS': INTERNATIONAL LAW
 INTERNATIONAL LAW

On the wing through Europe

The second Sexton cook book for the Sexton market (repetition of
 main entry, but not given as an initialism; best to arbitrarily
 abridge as:)
 THE 2D COOK BK./SEXTON MARKET

All's well that ends well, by. . .; ed., with notes, introduction, glos-
 sary, list of various readings, and selected criticism, by. . .
 (redaction/form)
 ALL'S WELL THAT ENDS WELL--ED.;/NOTES, INTRO., GLOS-
SARY, LIST/VARIOUS READINGS, + SEL. CRITICISM; BY. . .
 (traced as:)
 ALL'S WELL THAT ENDS WELL--ED. BY. . .

Thanksgiving plays for boys and girls, a collection of six. . . (form,
 occasion)
 THANKSGIVING PLAYS/GIRLS + BOYS; A COLLEC./6. . .

Final report of the. . . (corporate main entry repeated)
 FINAL REPORT

The young citizens league; helps and suggestions for organizing
 (form)
 THE YOUNG CITIZENS LEAGUE; HELPS. . .

Not a sparrow falls, a novel (form)
 NOT A SPARROW FALLS; A NOVEL

Manifesto of the Spanish nation to Europe
 MANIFESTO/SPANISH NATION/EUROPE

O C D, outline of lectures on civilian defense (alternative to abbreviation)

O C D, OUTLINE/LECTURES/CIVILIAN DEFENSE (Note need for spaces in abbrev. entries)

Sketching methods

The Swiss cross, a monthly magazine of popular science (title main entry, form/explanatory)

THE SWISS CROSS; A MONTHLY. . .

Shelley

The Germania of Tacitus, with ethnological dissertations and notes.

By R. G. Latham (supplementary/redaction)

THE GERMANIA;/ETHNOLOGICAL DISSERTATIONS + NOTES--BY
R. G. LATHAM

(traced as:)

THE GERMANIA--BY R. G. LATHAM

Carolina humor; sketches by. . . (form)

CAROLINA HUMOR; SKETCHES. . .

Brief memorials of an only daughter

BRIEF MEMORIALS/ONLY DAUGHTER

The distant princess

Tumminello theory of numerical roots (initialism)

THEORY/NUMERICAL ROOTS

The minority of Henry III.

THE MINORITY/HENRY 3D

The twin sisters; or, The effects of education: a novel; in a series of letters (title main entry, alternative, form)

THE TWIN SISTERS (THE EFFECTS/EDUC); A NOVEL. . .

THE EFFECTS/EDUC.; A NOVEL. . .

A history of philosophy, from Thales to the present time (scope)

A HISTORY/PHIL.;/THALES. . .

Under the crown, American history; selections from the Youth's companion (title main entry, explanatory, form)

UNDER THE CROWN--AMER. HIST.; SEL./ . .

The Unitarian; conducted by. . . (title main entry)

THE UNITARIAN; . . .

The application of statistical methods to the problems of psychophysics

THE APPLICATION/STATISTICAL METHODS/PROBLEMS/
PSYCHOPHYSICS

The geology of the country around Newton Abbey (scope)
 THE GEOLOGY/COUNTRY AROUND NEWTON ABBEY

Studies in the Asclepiadaceae
 STUDIES/ASCLEPIADACEAE

An essay on the antiquity of the Irish language, with a preface, proving
 Ireland to be the Thule of the ancients (supplementary)
 AN ESSAY/ANTIQU./IRISH LANGUAGE; WITH. . .

Hand book of the United States tariff, containing the tariff act of 1897
 rev. to July 1, 1902, with complete schedules of. . . (explanatory)
 HAND BK./ U. S. TARIFF; cont. . .

Deterioration of steels in the synthesis of ammonia
 DETERIORATION/STEELS/SYNTHESIS/AMMONIA

An integrated course of study in speech
 AN INTEGRATED COURSE/STUDY/SPEECH

The story of milk and how it came about
 THE STORY/MILK + HOW IT CAME ABOUT

Paintings of the sea
 PAINTINGS/SEA

An introduction to English grammar, on an analytical plan, adapted
 to the use of students in colleges. . . (form/opportunity)
 AN INTRODUCTION/ENG. GRAMMAR,/ANALYTICAL PLAN;
ADAPTED. . .

Laird & Lee's vest-pocket Webster pronouncing dictionary. . .
 (initialism for publ.)
 LAIRD + LEE'S VEST-POCKET WEBSTER PRONOUNCING DICT.

Anticlinal structure in parts of Cotton and Jefferson counties, Okla-
 homa (scope)
 ANTICLINAL STRUCTURE/PTS./COTTON + JEFFERSON COUN-
TIES, OKLA.

The analytical chemist's assistant: a manual of chemical analysis,
 both qualitative and quantitative of natural. . . (explanatory)
 THE ANALYTICAL CHEMIST'S ASSISTANT: A MANUAL. . .
 A MANUAL/CHEMICAL ANALYSIS, BOTH QUALITATIVE + QUANTI-
TATIVE. . .

The Greek romances in Elizabethan prose fiction
 THE GREEK ROMANCES/ELIZABETHAN PROSE FICTION

Memorials of James Wood, LL.D., J.P. of Grove House, Southport
 MEMORIALS/JAMES WOOD; LL.D. . .

The silence, thoughts on the silence, and on various other subjects
(explanatory)

THE SILENCE; THOUGHTS. . .

Punches, dies and tools for manufacturing in presses (scope)

PUNCHES, DIES + TOOLS--/MANUFACTURING IN PRESSES

Social life at the English universities in the eighteenth century (scope)

SOCIAL LIFE/ENG. UNIV./18TH CENTURY

The filing of such a sample of entries has not yet been undertaken, though its satisfactory use is obviously necessary before such a compression can be judged acceptable. However, the principles and desiderata enunciated in the preceding text, together with the Hines/Harris code, should produce the basis on which to attain such an acceptability. The two major problems, of course, are the filing of abbreviations with other similar-stemmed words (a problem encountered in the filing of L.C. subject headings, with their compulsory abbreviations, already), and of the filing of the virgule (=prepositional-phrase-introducers)—this last particularly in that the same symbol equivalated (in the recommendations) to a multiplicity of such prepositions, from a ("at". . .) almost to z ("with". . .).

INFORMATION RETRIEVAL PROJECTS IN THE BIO-MEDICAL LIBRARY, UNIVERSITY OF CALIFORNIA, LOS ANGELES

Louise Darling

In the area of information retrieval, the Biomedical Library is currently engaged in two projects: the UCLA MEDLARS Station and the Brain Information Service. This paper is limited to a condensed and non-technical consideration of their machine programs. The two projects are interrelated through the National Library of Medicine's Medical Literature Analysis and Retrieval System (MEDLARS) which must therefore be our starting point. At the outset it should be made clear that this is a report on work in which the author's role is purely administrative. The progress and accomplishments described are due to the efforts of the staff of the projects.

UCLA MEDLARS Station

MEDLARS was developed at the National Library of Medicine (NLM) during the years 1961-1963. It operated on a provisional basis in 1963 and in 1964 became the sole system for producing the Index Medicus, other recurrent bibliographies in special subject areas and individual-oriented demand searches. The file consists of citations in sequential order on magnetic tape with each citation carrying its own set of descriptors. For each record, the journal title and descriptors are in code, all else is in English. Searches are batched, but the entire file (or a specified portion by date) is searched for each request; i.e., as presently designed, the file is not open to random access. Around 40,000 citations were put into the system in 1963, but generally searches are begun with the 1964 store since the literature representation for 1963 is so incomplete. Counting from 1964, the file now contains over 350,000 citations and grows currently at a

Louise Darling is the Librarian of the Biomedical Library of the University of California at Los Angeles.

rate of approximately 14,000 citations per month. The descriptors comprise the Medical Subject Heading List, usually referred to under its acronym as MeSH. MeSH is a rigidly controlled vocabulary with hierarchical features. It and other components of the system will be considered later in relation to the UCLA Search Station. MEDLARS programs are written in machine language (Argus) for the Honeywell 800 Computer. The flow chart in Figure 1 gives a very general idea of the system as it operates at NLM.¹ The system is under continuous study and subject to frequent revision.

Recognizing that enormous demands would be placed on the system if it lived up to the intent of its design, NLM early developed the concept of reproducing its data tapes for use in installations across the country, thereby expanding manifold access to a, if not the principal part of the world's biomedical literature. Large scale decentralization presented two major problems. First, the programs required rewriting in a language which would make for easy adaptation to other computers comparable to the Honeywell 800; second, a mechanical conversion of data from the 3/4 inch tapes used by Honeywell to the 1/2 inch tape used by IBM and other equipment must be made. Fortunately, the means to do the latter at relatively low cost had already been developed.²

Plans for decentralization went forward. In November 1964, NLM awarded a contract to the University of California at Los Angeles to rewrite the programs in the COBOL language and, following this, to establish a pilot demonstration search center which would test the programs under actual operating conditions. Only programs for the retrieval subsystem were specified, of course, since neither input nor publications subsystems are a part of the decentralization plan. The renewal contract for March 1966-February 1967 specifies the Pacific Time Zone, Alaska, and Hawaii as the area the Station is to serve. A second contract was awarded in January 1965 to the University of Colorado Medical Center to operate a pilot search center using compatible equipment. Under the direction of Dr. Frank B. Rogers, this station successfully served UCMC from February to August; then under terms of a new contract, in September 1965, opened to the medical community in the Mountain Time Zone. Much experience valuable to the whole decentralization program has been gained from the pioneering work at Denver.³ Three additional stations have recently been authorized to serve the Northeast, the South, and the Midwest. (See Figure 2.) These will use the UCLA COBOL programs modified as needed.

Rewriting the MEDLARS programs at UCLA has been the responsibility of the Health Sciences Computing Facility (HSCF) (which is sponsored by the Division of Facilities and Resources of the National Institutes of Health through grant FR-3.) The Facility is a very busy center operating twenty-four hours a day, seven days a week. It

deals primarily with computations and data processing for scientific research conducted in the Center for the Health Sciences. Several off-campus users are served through remote inquiry stations. A staff of programmers is available for advice and assistance, but large projects using the facility usually employ additional programmers. The principal equipment complex is now an IBM 7040-7094 direct couple configuration. However the IBM 7040-7094 complex is scheduled to be replaced well before the end of the calendar year by an IBM 360 model 75 system. Implementation of the new system is now under way and users are preparing for necessary computer program modifications. (The IBM 7094 is a stand alone configuration but is almost always operated in a direct couple mode.) Its runs fall into four categories. In descending priority these are: Category 1, jobs of eight minutes or less; Category 2, fifteen minutes or less; Category 3, thirty minutes or less; and Category 4, over thirty minutes. The latter generally are run at night and on weekends. (Exceptions are made, of course, and the pattern itself adjusted as necessary to the overall workload in the Facility.) The Facility is thus geared for maximum computation time efficiency in line with needs of the great majority of its users. Those who do not fall into this group must adapt to it.

Because programmer recruiting is always a problem in Southern California, especially for fixed term projects, HSCF decided to subcontract for the MEDLARS conversion and at the same time to train a resident MEDLARS programmer who would carry on as part of the Search Station after the initial job was completed. The subcontract was awarded to the Planning Research Corporation (PRC) and has proved a satisfactory arrangement. Liaison with NLM has been effected through its Information Systems Division and via the rather extensive training program in MEDLARS indexing and searching techniques given four members of our staff who spent from three to five months each at NLM during 1965. The vital need for constant interchange between the parent system and satellite centers cannot be overemphasized, especially since file maintenance and updating are handled centrally.²

The NLM-UCLA contract stipulated that the system retain its ability to accept all forms of search request statements and to produce all standard report forms. However, because of the wide differences between the Honeywell 800 and the IBM 7094 in both hardware and software, the UCLA system was not restricted to specific guidelines of the NLM system, although the system logic in general was to be retained. The COBOL requirement has, in most respects, worked out satisfactorily, for not only is the language adapted to handling large quantities of MEDLARS type data, but it is a fast language for writing programs for systems of the MEDLARS type. During the implementation phase, no COBOL compiler nor COBOL

systems errors were found, although there were difficulties due to lack of IBM documentation. Use of COBOL compared with machine language has not lengthened searching time quite as much as some predicted, but neither is it proving as efficient as appeared to be indicated six months ago. The section most costly in machine time is the high speed or gross search submodule (HSS) which has been averaging two to two and one half hours for the 1964-1965 file. Since the HSS is based on retrieval of coded descriptors, not on English words, it would be feasible to redo the routines for it in machine language. Some installations might also wish to consider redoing some of the Log Search routines.⁴ (The search routine programs were rewritten in MAP during the summer of 1966 at UCLA.)

Problems in implementing the UCLA system have not been as numerous as might have been expected in view of the wide divergence between the UCLA and NLM equipment. There were however, two serious tape file problems: tape bad spot handling and character writing. The Honeywell 800 does not bypass bad spots while writing, but the IBM 7094 does bypass these spots in reading, thus losing data. The second problem arose in the tape conversion process. The Honeywell 800 writes only in multiples of eight characters (one computer word consisting of 48 bits in 8 bit units) while the IBM 7040, which at UCLA handles the tape for the IBM 7094, works in multiples of six characters (36 bits to the word).⁴ NLM, which handles the first step in the conversion in-house through a satellite of the Honeywell 800, is modifying the routines for this step; meanwhile UCLA uses the IBM 7094 which can handle eight character words instead of the IBM 7040 to complete the conversion. Other difficulties with tape reading continued for some time. Among the trouble causers was the closeness of the load point to the first record and the use of high density tape (800 bits per inch), apparently corrected by return to 556 density.* Another stubborn tape read problem, although not one connected with the conversion, appeared to be tape stretching on the rewind after High Speed Search, a condition which necessitated several readjustments of the tape drives.

A word needs to be added on transference. The UCLA programs can be fairly easily modified for computers of like size and configuration, providing a COBOL compiler is also available, but, to quote Frank Garvis, project manager for PRC, "The use of the COBOL language does not mean MEDLARS is now easily transferable, with only minor changes, on any computer configuration. To put the UCLA MEDLARS system on smaller computers, such as IBM 1401 and 1410, would require major and costly revisions to bring it within the confines of the normal small machine hardware and software configuration."⁴

*The rapid growth of the file caused us to try high density tape again in the summer of 1966, this time with success.

There have been a number of changes in the NLM system since the UCLA contract was awarded. The two of most immediate importance for station operation are the addition of subheadings to the MeSH as of January 1966, and program preparation for the inclusion of citations for monographs. The latter require a format different from that for journal citations. The routines for these additions and a program to annotate the master citation file to indicate UCLA current serial holdings, as is done in the Colorado Search Station, will be completed within a few weeks as an extension of the subcontract with PRC. There are several more additions to the UCLA system which the resident programmer has under way or will begin soon. Among these are the option for restricting the HSS to a selected portion of the master citation file, beginning with the most recent citations, a check for possible loss of data on the converted tapes, implementation of the option for listing citations under any desired headings used for the search, allowance for printing out any subset of a search or any combination of subsets, choice of specifying tracings for one subset and not for another, and increase in the number of terms which can be specified per search. (Main heading terms or descriptors are now limited to 100, for example.)

With the transfer to the IBM 360 system, computer time sharing and on-line display devices to assist in both file maintenance and searching techniques should be considered. It has also been suggested that the master citation file might be inverted so that each descriptor carries a list of its citations instead of the reverse arrangement which now prevails. Random access would reduce computer time, but such a change would also enlarge the file and necessitate rather complete reprogramming.⁴ NLM is at present planning a study of redesign of the system. Obviously, there is an almost unlimited horizon for study and for change.

Now let us turn to a brief scrutiny of the operation pattern we are attempting to follow at UCLA. It corresponds in general with that of NLM. Figures 2-12 trace the journey of a search request. The requests on the Station's form (Figure 3) are sent by mail or brought in person to the Search Station where they are assigned for formulation to one of the search specialists. The most important part of the formulation is the selection of descriptors or main subject headings from the printed MeSH and from the provisional headings sent us by NLM. For use in the computer search, MeSH is put in alphabetical order on tape and in this form known as the Medical Dictionary Tape (MDT). This tape also contains a tally of the number of times each heading has been used. NLM sends the Station an updated MDT every month. There is not time to discuss MeSH in any detail, but it should be remembered that descriptors are given in both alphabetical and category position. Some categories have been exploded for machine searching to form tree structures. A tree

structured category is one where terms have been hierarchically ordered from generic to specific, using four levels of classification. The tree can be entered on any of the four levels and searched for this level and one or more levels under it. The system allows for 127 tree structures, many of which are in the process of development.⁵ Citations are given an average of eight to ten descriptors compared with an average of two or three in the printed Index Medicus.

Searching is done through the logical operators and (*) or (+), and not (-). An important feature of the system is the provision for as many as three search expressions or subsearches, A being the most general, B less general, and C the most specific. This makes it possible to test retrieval hypotheses, i.e., to use both broad and narrow approaches simultaneously. The relationship between the subsearches is used in selecting the most significant elements for the gross or high speed search, for what satisfies the A search obviously should cover B and C which are refinements of A.⁵ All searches are not however, required to have three, or even two expressions. "A significant element is a term or search element with the lowest index usage in a search and therefore of highest retrieval value. For example, in a search on thalidomide poisoning in rabbits, the significant element is thalidomide. In a search with Or'd expressions there may be as many significant elements as there are combinations. Thus in a search on new uses of dimethyl sulfoxide or procaine, both drugs are significant elements. On the other hand, for recent work on kidney transplantation in dogs or monkeys there is just the one significant element, kidney transplantation."²

Figure 4 shows the side of the formulation sheet on which all the elements for the search have been listed, including unions of terms. This is the Request Element Definition List or RED List. Figure 5 shows the other side of the formulation record where the essential information about the requestor and the request are given in the upper section and the search equation in the lower. IBM columns are indicated and type of IBM card to be punched, e.g., header cards (02), RED List cards (03), elements and operators for subsearches (04-06), and report generator cards (07).

A decklet of punched cards for the search is prepared from the formulation sheet, the decklets batched, then sent through preliminary Edit (see Figure 6). This is done as an additional safeguard to proof subject heading cards and to obtain a tally of the number of citations which the significant elements can be expected to retrieve. Tallies can also be taken manually from the print-out of the MDT and added, but it is more convenient to obtain them through the Edit Program which only takes 3 to 4 minutes to run and may save hours of time by detecting errors such as off-center keypunching and misspelling in advance of the main run. The tally is essential for

estimating the time it will take to run the batch. This information is required for all jobs submitted to the Computing Facility. After preliminary Edit some components of the batch may be removed for separate running if the tally has shown a very high retrieval count. The batch or batches then go to HSCF with the proper control cards for the main run (see Figure 7). Note the Intermediate Retrieved Citation File tape. If there is insufficient time for the entire search, or if the search blows up in Logical Search, the citations retrieved from HSS may be stored off-line on tape and run again later through Logical Search. In Logical Search the citation output from HSS is searched against the search formulations in such a way that only the citations which logically satisfy the requests are selected. An intermediate tape with the results of Log Search may also be saved for later processing through Report Generator if there is need to do so. Processing of the batch may also be stopped at this point to obtain a statistical report on number of citations retrieved per search so that searchers may check on their formulations and rewrite them if the count indicates trouble. Figures 8-10 illustrate formats in which the completed search bibliographies may be obtained; figures 11-12 make up the request and form for user appraisal of the search.

Our experience thus far has shown batching to be of critical importance. The MEDLARS file is a large one, growing rapidly larger. When it is searched with the COBOL programs, machine time used increases markedly. Thus it is imperative to make optimum use of available time. The maximum batch allowed by the UCLA system appears to be more than adequate—something in the realm of 100 searches with an average of 50 machine words each, the limit being set by the request table for the coded search expression. This table provides locations for a total of 5,250 machine words. The present limit can be raised, but it is important to keep in mind that as batch size increases, the difference between the number of citations searched in the HSS and the number searched in Log Search lessens, thereby cutting the advantage gained from the gross search feature of the system.² The possibility of tape runover in HSS with large batches, which has been a problem in the Colorado station, causes no concern in the UCLA station because disk is used in place of tape. (There was a problem when intermediate tape files were stored off-line, but this has been corrected by program modification.) The problem in our case is running over estimated time, but the end result, inability to complete the search, is the same.

To date we have generally used batches of ten to fifteen searches, but we have gone as high as thirty. Processing a batch through the entire file generally takes from four and one half to five hours. The length of time for HSS depends primarily on the size of the file, not the size of the batch. The only batch factor of importance is the proportion of significant author elements to significant descriptor

headings. Because authors are in English, and more characters are involved, searching them takes longer. It has consistently been in the range of 1,400 to 2,000 citations per minute, with the average usually in the lower end of the range. Because of the relatively limited variation in time, an experiment was carried on with the 30 batch search. The Intermediate Retrieved Citation File generated in HSS was used as a selected CCF. The batch was broken into a group of smaller batches, each of which was run in HSS against the selected CCF, then, as time was available, through Log Search and Report Generator where time used is a direct expression of the number of citations retrieved to satisfy the search request statement. This procedure made it possible to submit some computer priority 2 and 3 jobs and thus to obtain results much earlier than is possible with priority 4 jobs. In addition to this bonus, batch segmentation makes MEDLARS jobs less liable to blow up because the probability of encountering errors is smaller. Moreover, if there is a blow-up, the penalty in time lost is not so great, although the total amount of machine time used by this type of batching is greater than the more direct procedure providing the latter does not run into tape read errors or other trouble.

These may very well prove important considerations for routine operations; they are of particular value during the developmental period when search failure and search completion are in decidedly unfavorable balance. A multiplicity of problems is normal for any volume testing period, but in our case they have been compounded by the length of the MEDLARS file and the sophistication of the computer configuration. On the first score, for example, MEDLARS already has well over 300,000 records and the maximum acceptable margin of error on HSCF's IBM 7040-7094 system is one error in 300,000 records, or a mean free error time of forty minutes. On the second score, running time varies for a MEDLARS job in Edit or Report Generator, where input-output is an important factor, with the nature of other jobs going through the IBM 7040. The Edit routine, for example, has varied from 3 to 5 minutes on identical search batches at different times. Another problem is that the HSCF computer system, fails to skip over tape errors and instead throws out the search. This procedure is an advantage for mathematical computations since an error anywhere along the way invalidates the results, but in library-type manipulations it has a high cost in lost time. If this were changed, the error could be identified by manual examination in most cases and the printout corrected manually also. In any case, all the searches could then be rerun as part of another batch.

The reprogramming called for in the first contract was largely completed by November 1965. The first testing of the system as a whole was done by running parallel searches with a group which had already been done at NLM and the Colorado station against the UCLA

programs. Retrieval proved to be identical. We plan to continue comparisons periodically and to ask NLM to run a sampling of UCLA searches.

As of April 15, 1966, the Station has accepted 93 original searches of which 32 have been completely processed and returned to the requestors. Requestor reaction, searcher check, and evaluation of results have all been reassuring. The other searches have all been formulated and are in various stages of machine search, most of them at Report Generator whose programs have been the last to receive a thorough working over. Once the Station is operating routinely and well, requests should be processed in two weeks, with emergency service in one, providing machine time is available and search volume within reason, but we cannot yet predict how long it will be before this happy state of affairs is at hand.

Statistics are collected from all batches on the number of searches in the batch, the machine time for Edit, HSS, Log Search and Report Generator, the number of citations searched, retrieved and printed, the number of significant elements, and the number of citations printed with tracings or groups. The latter and the number of citations printed are also kept for individual searches as is the processing time from arrival of request to return of the MEDLARS produced bibliography. However, the small volume handled thus far does not lend itself to support of any generalizations, and the stop and repeat manner in which searching has proceeded in the test period makes figures for total length of time to process requests meaningless.

Timing comparisons between the COBOL and the NLM machine programs are of special interest, but there are as yet no reliable data on which to base them. There is some indication that the NLM programs may be nearly three times faster in HSS and twice as fast in Log Search. Transfer to the IBM 360 System should improve present timing considerably, perhaps by a factor of one, but it is now obvious that for stations running on IBM 7090 or 7040 series equipment, the search routines must be returned to machine language.

Change and experiment are certain to be the most constant characteristics of the UCLA Station for the next several months, in fact for the next year, but our experience to date should greatly facilitate the tooling up period for the three new stations now in process of organizing.

Brain Information Service

The UCLA MEDLARS Station has also made a substantial contribution to the Brain Information Service (BIS) which was established at about the same time in 1964 under contract with the National

Institute of Neurological Diseases and Blindness. Its purpose is to provide information in the basic brain sciences rapidly, accurately, and completely to investigators and teachers in this field. The subject area of the Service includes neuroanatomy, neurophysiology, neuroendocrinology, neurochemistry, neuropharmacology, and non-clinical aspects of neuropathology. The literature of diagnosis and treatment of neurological diseases is not within its scope.⁶

For convenience, the Service is referred to in this paper as a Biomedical Library project, but it is only partly so as the organization chart in Figure 13 indicates. The Scientific Unit is responsible for conferences and workshops, stimulation of the writing of critical reviews, a study of the information gathering habits of scientists working in the brain sciences, and for formation of the glossary which is the vocabulary used by the computer in searching. The Bibliographic Unit is responsible for input for the computer store and for the search and current alerting service, both manual and automated. At the present time only manual service is available, and it is limited largely to members of the UCLA Brain Research Institute who serve as a users test group. On request several investigators in other California institutions and in other parts of the United States, as well as a few abroad, have been added to the group. Toward the end of 1966 when the computer programs will have been thoroughly tested and the glossary more adequately developed, the Service will be opened on a national scale. The programs are all written except for the Report Generator and Publication Modules which are due next month and a glossary maintenance program which will not be ready for some months yet.

BIS uses the same equipment in the Health Sciences Computing Facility as does MEDLARS and also has handled its initial programming through subcontracts with the Planning Research Corporation. Every effort has been made to secure compatibility between the two projects where this is practical and to keep personnel on one informed of the operations of the other and vice versa. The BIS programs have been written in COBOL, and their modifications for the new IBM 360 System are being planned jointly with those for the MEDLARS Station with the help of the Computing Facility.

In addition to MEDLARS-type demand search routines, BIS programs fall into two groups: input and glossary. Their end products produce a master citation file which may be searched with the same type of formulation as is employed for MEDLARS although descriptors are drawn from a quite different type of vocabulary. The vocabulary will be described along with the glossary programs.

The sources of input for the BIS data base are shown in Figure 14. The first is the magnetic tape which brings a cut of the UCLA MEDLARS file drawn off by recurrent demand search, beginning with the 1964/65 file, monthly thereafter. Next are citations on punched

cards from the retrospective literature which includes a selection from the Cumulated Index Medicus, 1960-1963, and from bibliographies compiled manually by subject area specialists in the Bibliographic Unit. A selection from other sources such as Biological Abstracts and the Current List of Medical Literature will be added later as time permits. Citations from a small core of the most important current journals, which are required for the system as soon as published, as well as a group of biological journals not included in MEDLARS, are indexed by the area specialists to form a third type of input. Current symposia, conference proceedings, and other multi-author volumes indexed by the area specialists form a fourth input. Overlapping between citations from the MEDLARS system and other input sources is deleted by the merge program from which the BIS Compressed Citation File is generated. Figure 15 gives an estimate of the size of representation of the various types of input in the file. A very important component not shown is the backlog of retrospective citations still to be keypunched from the Cumulated Index Medicus 1960-1963, which has been only 60 to 65 percent completed, and from other bibliographies. Information on research in progress in some fields is being received from the Science Information Exchange, but plans have not yet been developed for incorporation of this information into the computerized system.

The citation entry record (see Figure 16 a-c) is arranged for easy transference to keypunch cards. It was designed to be as flexible as practical and to include space for certain features which would be desirable to add at some future time such as indexing values for terms and relationships for terms. (See Figure 16a, Format, Columns 73-80.) Incidentally, the word "Link" as used on this form means simply that cards and columns are linked in the case of run-over in number needed (see Figure 16b). Role indication columns at present are used only to show the source of the term.

The "10" cards give all the usual bibliographical information about the citation, name of indexer, sequence number for citation, and source of citation. The "20" cards are subject term cards. Abstracts, when included, go on "30" cards, and authors and editors go on "40" cards. Abstracts may consist of editorially selected portions of the text of the article. All authors, regardless of number, are punched and, when the citation is retrieved, printed out. Titles in English, both original and translated, go on "50" cards, and titles in the vernacular or transliterated go on "60" cards. There are as yet no "70" cards. The "80" cards are for the senior author's address, and the "90" cards are for the names of the scientists who have requested information on the subject of the article. ("90" card information is gathered for a small separate program which is part of a current alerting service.) Sort criteria are journal title code, volume number, year, starting page, and senior author. All five must be

punched in order to add, delete, or update a citation because these are also the merge criteria. In the case of chapters in multi-author volumes, (BK) is used for journal title code and zeros wherever else something non-existent in the citation must be supplied.

Programs convert the various inputs to the standard tape format for which the citation entry record was designed, merge identical citations into one record which contains all information from all sources, and extract words from titles for mechanical indexing of citations. (See Figure 17.) The extracted title words and all other descriptors attached to the citation are then matched against the glossary tape. If the words are recognized as indexable, they are added to the citation as coded tags. If the words prove to be synonyms, the terms with which they are synonymous are substituted and assigned to the citation, but if they are found to be non-indexable, no further action is taken. At present we are entering only English words as potential subject headings, but computer routines allow any Roman alphabet language to be used. Words not found in the glossary are printed out as warnings to the glossary editor who must then make the decision on the status to be given the new word in the glossary: indexable as subject heading, indexable as synonym, or non-indexable. Once this is done, a second run can be made for the purpose of attaching the new terms to the citation. The updated master citation file is then ready for searching by MEDLARS strategies.

Like citations in the master file, terms in the glossary are drawn from MeSH, whose descriptors are appended to the citations from MEDLARS, from the subject headings used in Index Medicus 1960-1963, from words assigned by BIS area specialists who double as indexers, and from words in titles. Indexers assign terms from the text of the article, generally without reference to the BIS glossary. There is no limit to the number of words in a term assigned by indexers except that potential descriptors in the glossary are limited to forty-eight characters and/or spaces.⁷

As previously noted, all citations entering the system from any source are mechanically indexed from words occurring in their titles. Terms are also drawn mechanically from abstracts introduced on BIS citation entry records. These terms consist of every single word and every two words in sequence except for a list of 127 words which are inadmissible, such as A, ALSO, BEFORE, IN, IS, OF, STUDIES, THE, TO, UPON, USING, WHEN, and WITH. These are mechanically rejected on examination by the programs. In the title "Electroencephalographic studies in experimental pulmonary oedema" the potential terms are ELECTROENCEPHALOGRAPHIC, EXPERIMENTAL, PULMONARY, OEDEMA, EXPERIMENTAL PULMONARY, and PULMONARY OEDEMA. Routines do not assign a single word descriptor to an article when that word also occurs in a two-word sequence, thus, PULMONARY OEDEMA but not PULMONARY alone or OEDEMA.

If this restriction proves inappropriate, it can, of course, be changed later. In the mechanical indexing program, more than two terms in sequence are inadmissible. Thus for CIRCLE OF WILLIS, only the single words CIRCLE and WILLIS will be picked up.

The BIS glossary is an open-ended vocabulary based on natural language. Control is provided by the glossary editor who is responsible not only for decisions on which terms will be added as indexable and which declared non-indexable, but also relates each indexable term to its generically more general and more specific terms.

Procedures in glossary formation are illustrated in Figures 18-20. Figure 18 is a printout from the alphabetical section of the glossary (which in its computerized form bears the unfortunate acronym of BISH for Brain Information Subject Headings). The numbers in the first column at the left are codes for all indexable terms whether subject headings or synonyms. Non-indexable terms are indicated by a zero. Terms with the same numbers make up a synonym set. In the next column, the figure 1 indicates both a subject heading and a non-indexable term, 2 a synonym. Under "tag source" is the letter I used for an indexable MeSH or non-MeSH heading, M for a non-indexable MeSH term and N for a non-indexable non-MeSH term. All MeSH terms will become indexable as soon as citations using them enter the computerized system. They are entered in the glossary as unindexable now for convenience of the editor who will not place them in the hierarchy until they appear for the first time on citations. The "frequency column" is used to show the number of times a synonym set has been assigned to a citation. In this case all are zeros because only an initial small group of citations had been entered into the system at the time of the printout, and the second run for the glossary update had not yet been made. The "specificity column" will not be used until later. Then it will give an index value for the term, probably 1, 2, or 3 representing general, less general, and least general; or perhaps just 1 and 2, general and specific. The index value, in turn, will be used to review the specificity with which citations have been indexed. All citations whose subject tags total less than a given value will require further indexing. Details are yet to be planned. In reviewing, for example, it may prove more effective to omit the value number for general from the total count for the citation since the purpose of the review is to increase the specificity. Since many citations will enter the system with subject terms drawn only from mechanical extraction from titles, a review of indexing effectiveness is essential. The "main heading column" lists all the descriptors in alphabetical order. Note in this column the terms coded 36120. These are part of a synonym set, with the term electroencephalography indicated as the preferred heading, the other seven terms as synonyms. Members of a synonym set may, and usually are, separated in the alphabetical listing. In the Numerical BISH (see Figure 19) into

which the Alphabetical BISH is sorted for the purpose of attaching subject headings in code to citations, the order is numerical, and all synonyms are brought together without regard to alphabetical order. Thus the complete synonym set for electroencephalography can now be seen to consist of fifteen members. More may be added later as the glossary expands.

Figure 20 shows warnings for two citations from the system to the editor. It presents him with eight terms on which he must make decisions for the glossary. The other fourteen terms are already in the glossary as either indexable or non-indexable terms.

Figure 21 illustrates the hierarchical features of the glossary. At present the hierarchy is maintained clerically and used chiefly to locate synonym sets in the glossary when new terms are under consideration for inclusion in a given set. The order is not necessarily one which is appropriate for explosion of terms in a search; in many cases the editor will later rearrange it for more effective explosion of general terms. Moreover, there are as yet no means for mechanically exploding terms, although a program is planned to provide this capability and to make it possible to enter for retrieval at any level rather than to follow the tree structure pattern in MEDLARS which allows only four levels of specificity.⁷

A glossary maintenance program which will eliminate a good part of the clerical upkeep for building the glossary and will provide a more rigorous check for errors is expected to be ready soon for modification by BIS. The program will replace the checking for locations in the hierarchy and the adding of synonyms manually as well as the manual insertion and positioning of preferred new subject headings. This program has been developed by IBM for Dr. Baldwin Lamson's research at UCLA on recording and retrieving clinical pathological records in natural language.

The routines used for development of the glossary will obviously increase its size very rapidly as the first citations are entered into the system. They will also produce a very large indexing vocabulary whose rate of increase per entered citation will, however, eventually be quite low because most of the potential descriptors will already be in the glossary. As of March 1, 1966, the glossary contained over 6,000 indexable terms in about 5,000 synonym sets and 1,100 non-indexable terms. It does not appear that the size of the glossary will have a significant effect on machine time required to operate the system. Initial experience is that approximately twenty minutes, including time for input from card to tape, are required to take the terms from a batch of 200 citations. The major segment of time for this routine is used in sorting single and double terms into alphabetical sequence after the trivial words are removed. A maximum of 600 synonyms can be added at one time, but there is no limit on the number of preferred subject headings which may go in together.⁷

The size of the glossary and its continuous growth obviously complicate glossary maintenance and use, but the glossary has advantages which we hope will more than compensate for its disadvantages. One of these is mechanical indexing from titles, a method which is very cheap indeed for it involves no indexer time at all and no extra keypunching. On the other hand, it does not generally provide depth adequate for needs of a service such as the BIS. However, applied to abstracts this method is likely to prove very valuable and involves little more than the additional cost of keypunching the abstracts at the time they are added to the citation entry record. The cost of uniformly adding abstracts to the system is, of course, another thing apart from their use for indexing. In any case, it has never been the plan to rely entirely or even heavily on mechanical indexing, but to use it as a method to build up a file quickly and passably well. The screening for specificity of indexing referred to earlier will make it possible to bring all citations up to about the same level of indexing eventually. Meanwhile the mechanical indexing serves as a very useful supplement to what is brought in from MEDLARS indexing and BIS staff indexing.

Another advantage is that human indexing can be done less laboriously with this type of glossary than with a rigidly controlled vocabulary. Since words from texts are what BIS indexers are asked to use when possible, the indexer need not look up terms but can substitute underlined text for filled out indexing sheets. Further, intelligent but unskilled indexers can perhaps be utilized for some of the indexing because it is generally evident from the syntax which descriptors should be chosen from the text, even if the indexer does not know the meaning of a particular word and does not look it up. For example, the term GUANAZALO (an antineoplastic purine analogue) will be placed in its synonym set under the preferred heading 8-AZAGUANINE by the editor the first time it occurs and mechanically thereafter. The intellectual effort required to relate the two terms is thus expended only once.⁷

There is expectation, too, at least among the scientific staff, that consistency of indexing will be higher with this open type of vocabulary than with a controlled vocabulary because it is assumed that "authors are more consistent among themselves in their use of technical terms than indexers are in relating authors' technical terms to an approved vocabulary."⁷ Moreover, indexing with this type of vocabulary should not only, on the whole, produce greater specificity but should also reflect new subjects and new terms in the literature as soon as citations are added to the data base.

MEDLARS is a general service for the biomedical community and thus a first line support for the highly specialized Brain Information Service. BIS will handle more kinds of information, but its total file will be far smaller. While the MEDLARS Station looks to

NLM for future direction and development, BIS must look to itself but take full advantage of whatever it can usefully incorporate from MEDLARS and from the other neurological information centers with which it works in close cooperation.

Grateful acknowledgment is made to Mr. Robert Braude, Head Searcher for the UCLA MEDLARS Station, Mrs. Pat Walter, Head of the Bibliographic Unit of BIS, and Dr. Peter Amacher, Editor of the Glossary, for advice and information for this paper.

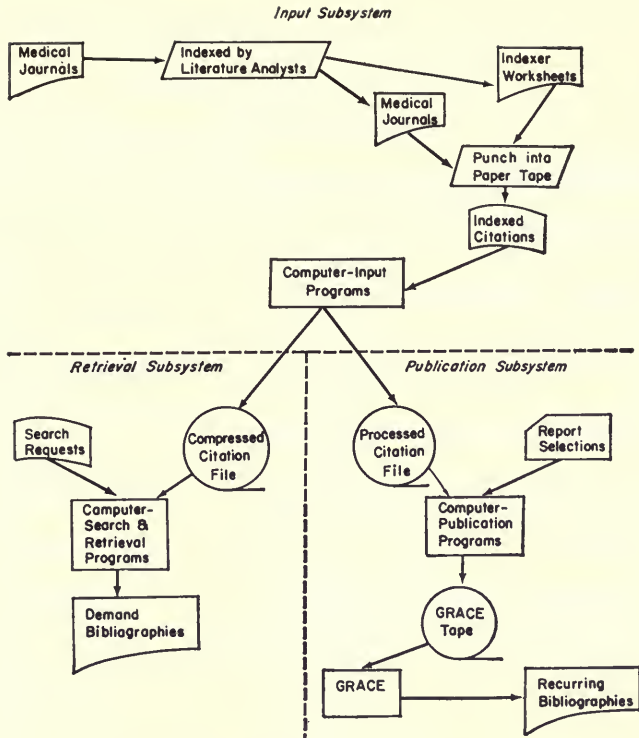


Figure 1
MEDLARS Flow Chart Reprinted from *Science*, 148:768,
May 7, 1966

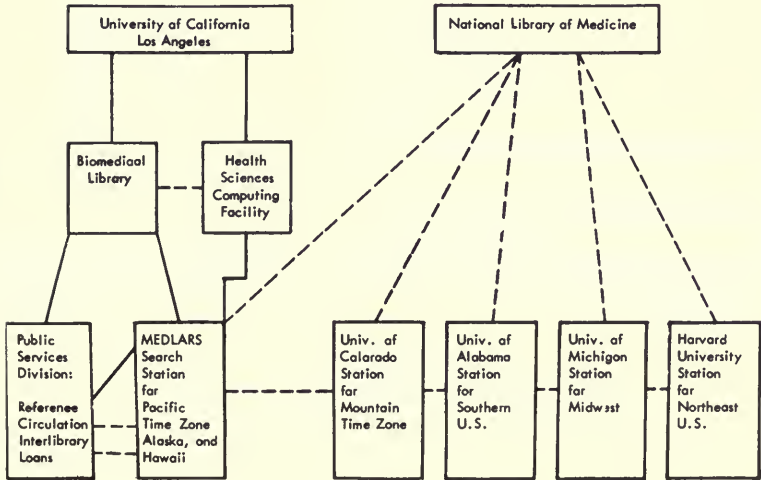


Figure 2
Organization of UCLA Search Station and its Relation to other Stations and the National Library of Medicine.

University of California, Los Angeles
Biomedical Library
MEDLARS Search Station

MEDLARS SEARCH REQUEST		Date
1. Individual who will actually use the bibliography		
John Q. Requestor, M. D.		
Title		
Professor of Medicine		
Organization		
Dept. of Medicine		
Address	Phone Number	
UCLA		
2. Request submitted by (if different from above):		
3. Detailed statement of requirements (Please be as specific as possible as to purpose, definitions, limitations, etc.)		
<u>XXXX</u> Human <u> </u> Animal <u> </u> Normal <u>XXXX</u> Pathological Bibliography on treatment of diabetes in old age. I am particularly interested in any differences between modified and unmodified preparations. <u>Age Range</u> <u> </u> to 1 <u> </u> 1-23 months <u> </u> 2-5 years <u> </u> 6-12 years <u> </u> 13-18 years <u> </u> 19-44 years <u> </u> 45-64 years <u>XX</u> 65- years		
4. A relevant citation(s) if known		
5. Limit Languages To: Accept all <u>XXX</u>		
English Foreign Specify:		
6. Print Specifications	7. Deadline:	8. Number of citations expected: Number of citations desired:
<u>XXX</u> 3" x 5" Cards <u> </u> Paper		

Figure 3
Search Request Form, UCLA MEDLARS Station.

DEMAND SEARCH FORMULATION RECORD

Request number:

1 6

ID	Seq	Elem Symb	Expl	Expl Level	Category Number	Element Name	Tally
7	8	9	10	11	12	13	14
03							
		M1				Hallucinogens	
		M2				Adrenochrome	
		M3				Cannabis	
		M4				Lysergic Acid Diethylamide	
		M5				Mescaline	
		M6				Perception	
		M7				Body Image	
		M8				Equilibrium	
		M9				Hearing	
		M10				Illusions	
		M11				Motion perception	
		M12				Sensation	
		M13				Sensory Deprivation	
		M14				Smell	
		M15				Space Perception	
		M16				Taste	
		M17				Time Perception	
		M18				Touch	
		M19				Visual Perception	
		M20				Pitch Discrimination	
		M21				Appetite	
		M22				Hunger	
		M23				Hyperesthesia	
		M24				Hypesthesia	
		M25				Kinesthesia	
		M26				Pain	
		M27				Paresthesia	
		M28				Temperature Sense	
		M29				Thirst	
		M30				Depth Perception	
		M31				Form Perception	
		M32				Orientation	
		M33				Size Perception	
		M34				Afterimage	
		M35				Color Perception	
		M36				Eye movements	
		M37				Figural Aftereffect	
		M38				Flicker Fusion	
		M39				Color Perception Tests	
		M40				Weight Perception	
		M41				Vigilance	
		U1				M1 - M5	
		U2				M6 - M41	

Figure 4
Demand Search Formulation Record

DEMAND SEARCH FORMULATION RECORD
MEDLARS SEARCH STATION,
BIOMEDICAL LIBRARY, UCLA

Request number:
1 6

Formulation time:
Date:

ID 7 02	Seq 8 9	Subsrch 10 11	12 Information Concerning Request 80											
			Name Title Address Effect of hallucinogens on perception with special reference to LSD.											
ID 7	Seq 8 9	10	11	Elements and Operators 80 Tally										
04 05	01 01	U 1 * U 2 M 4												
ID 7 07	Seq 8 9	Subsrch 10 01	Farm 12	Farmat 13	No. Cits 14 18	Trcngs 19	Sort Keys 20 25	Group Headings						
			P	A		X	SAJ	TZ 26	G 27	H 28	P 29	Y 30		

Figure 5
Verso of Demand Search Formulation Record showing the Request Element
Definition List.

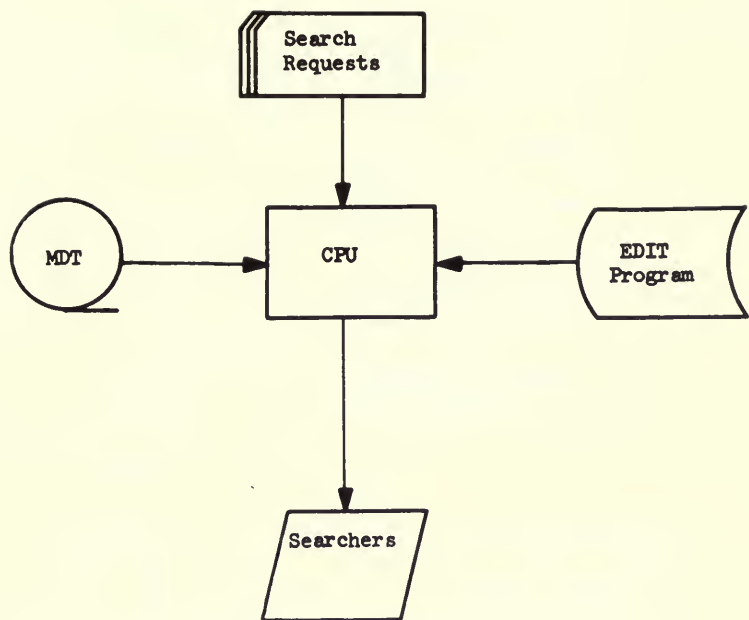


Figure 6
Flow chart for Preliminary Edit Program used in UCLA MEDLARS
Search Station.

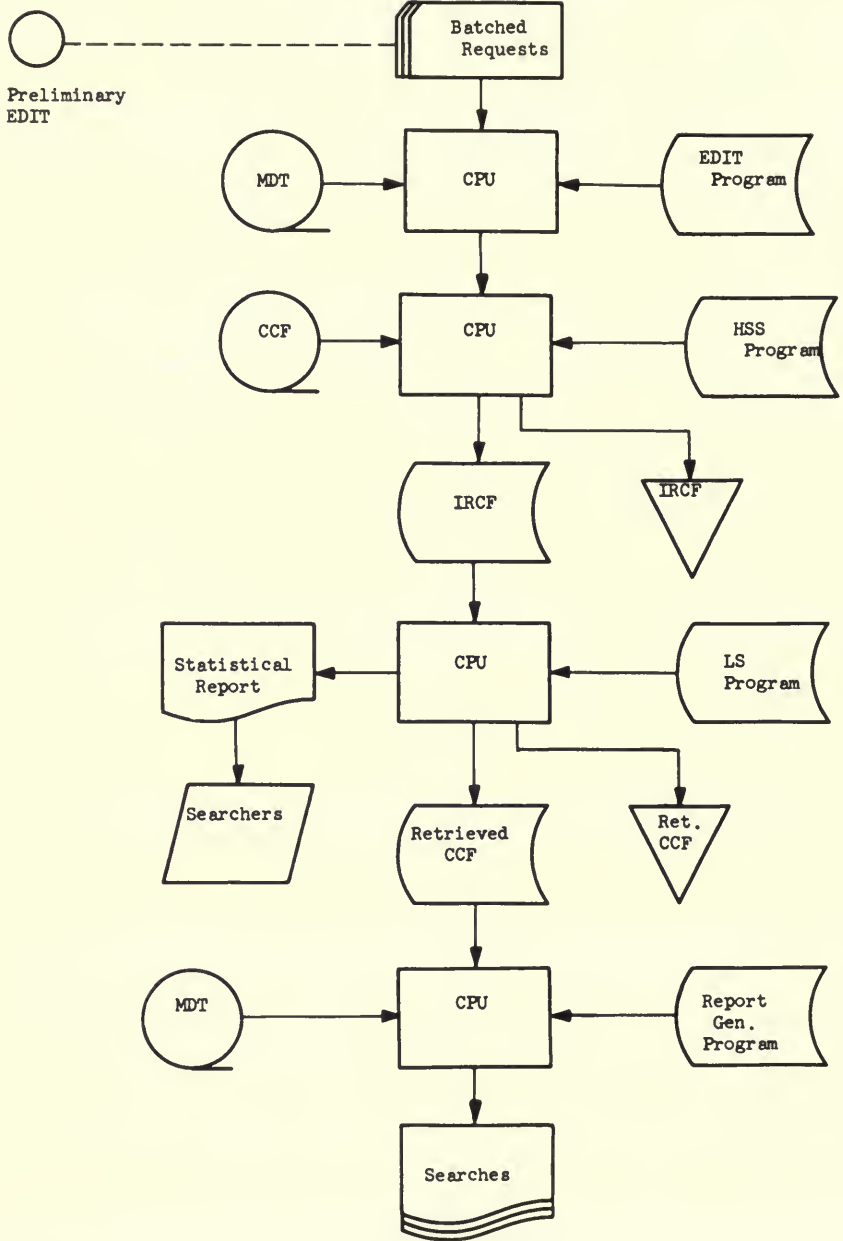


Figure 7

Flow Chart showing Search and Report routines, UCLA MEDLARS Station.

- ACTA PATH MICROBIO SCAND
84:50-4, 1965
LINDGREN I, ELDMAR E
A CYTOLOGICAL STUDY OF THE SEX CHROMATIN
IN CARCINOMA MAMMAE.
BARR BODIES, *BREAST NEOPLASMS, CYTOLOGY
GERIATRICS, *SEX CHROMATIN
- BRUNN BEITA KLIN CHIR
209:376-9, NOV 64
GLASER A, REISING R
THE SIGNIFICANCE OF NUCLEAR MORPHOLOGICAL
SEX DETERMINATION IN HORMONAL THERAPY OF
BREAST CANCER
*BREAST NEOPLASMS, *HORMONES, *NEOPLASM
THERAPY, *SEX CHROMATIN
- BULL SOC INV CHIR
23:17-23, JAN-FEB 64
BARAONAY G, NOMUS ZB
SEX CHROMATIN INVESTIGATION IN FEMALE
BREAST CARCINOMA
*BREAST NEOPLASMS, NEOPLASM ETIOLOGY,
*SEX CHROMATIN
- CHRONIK
98:318-8, JUL 64
REUELE M
ON THE PROBLEM OF SEX CHROMATIN IN MAMMARY
CARCINOMA
*BREAST NEOPLASMS, MASTECTOMY, MORTALITY
*SEX CHROMATIN
- DEUTSCH MED WISCHR
100:537-42, V APR 65
GROPP H, WOLF U, PERA P
SEX CHROMATIN AND CHROMOSOME STATUS IN
BREAST CANCER
*BREAST NEOPLASMS, *CHROMOSOME
ABNORMALITIES, PATHOLOGY, *SEX CHROMATIN
- DEUTSCH MED WISCHR
89:1215-7, 19 JUN 64
GROSS Y, HANNINGEN W, TREBBIN N, ET AL
ON THE SIGNIFICANCE OF BARR'S CELL NUCLEUS
BODIES IN FEMALE BREAST CANCER
*BREAST NEOPLASMS, *CELL NUCLEUS, *SEX
CHROMATIN
- GERMANSCH FRAUENHEILK
25:481-8, JUN 65
HECHTMANN U, POPP L, UHLMANN G
ON THE INCIDENCE OF BARR'S NUCLEAR SEX
CHROMATIN BODIES IN BREAST CARCINOMA
*BREAST NEOPLASMS, *CELL DIVISION,
GERIATRICS, *SEX CHROMATIN

MEDLARS SEARCH STATION-UCLA

Figure 8
Journal Format, MEDLARS Demand Search Bibliography

- SCHIFFINO A, GRANA E
EXPERIMENTAL CONTRIBUTION TO THE STUDY OF FETAL MALFORMATIONS CAUSED BY
DRUGS
PARMA (PARAF), 18:425-9, AUG 63
*ABNORMALITIES, EXPERIMENTAL LAB STUDY, HABITS, *THALIDOMIDE,
TOXICOLOGIC REPORT
- STAEHMLER N, HELM F, KIEL M
CONGENITAL MALFORMATIONS OF SKELETON IN RABBITS
MED EXP (BASEL), 10:22-6, 1964
*ABNORMALITIES, DRUG-INDUCED, BONE DISEASES, EXPERIMENTAL LAB STUDY,
*HYPOTHYROIDISM, PREGNANCY, *PREGNANCY, ANIMAL, RABBITS,
*THALIDOMIDE, TOXICOLOGIC REPORT
- STAPLES RE, HOLTKAMP OE
EFFECTS OF PARENTAL THALIDOMIDE TREATMENT ON GESTATION AND FETAL
DEVELOPMENT
EXP MOLEC PATH, 2:81-106, DEC 63
*ABNORMALITIES, DRUG-INDUCED, EXPERIMENTAL LAB STUDY, *MATERNAL-FETAL
EXCHANGE, PREGNANCY, PREGNANCY, ANIMAL, RABBITS, RATS, RIBOPHYLLIN,
*THALIDOMIDE, TOXICOLOGIC REPORT
- THOMAS E, PEARSE AG
THE SOLITARY ACTIVE CELLS. HISTOCHEMICAL DEMONSTRATION OF DAMAGE-RESISTANT
NERVE CELLS WITH A TPN-DIAPHORASE REACTION.
ACTA NEUROPATH (BERLIN), 3:238-49, 2 JAN 64
*BRAIN, *CARBON MONOXIDE POISONING, *CARBON TETRACHLORIDE POISONING,
*CENTRAL NERVOUS SYSTEM, *CENZYMES, EXPERIMENTAL LAB STUDY, FISH,
*HARTERS, *HYPOTHYROIDISM, *HYPOTHYROIDISM, RATS, RIBOPHYLLIN, *NEUROCHEMISTRY,
*NEURONS, *OXIDOREDUCTASES, *PSYCHOSURGERY, *PYRIDINES, RABBITS,
RATS, REPTILES, *SULFANTHIMIDE, *TETRAZOLINE SALTS, *THALIDOMIDE,
TOXICOLOGIC REPORT
- TUCHMANN-DUPLESSIS H, MERCIER-PAROT L
APPROPS OF TERATOGENIC TESTS. SPONTANEOUS MALFORMATIONS IN RABBITS
C R SOC BIOL (PARIS), 138:666-70, 1964
*ABNORMALITIES, *ABNORMALITIES, DRUG-INDUCED, EXPERIMENTAL LAB STUDY,
RABBITS, *THALIDOMIDE, TOXICOLOGIC REPORT
- TUCHMANN-DUPLESSIS H, MERCIER-PAROT L
REPERCUSSIONS OF NEUROLEPTIC AND ANTITUMOR AGENTS ON PRENATAL DEVELOPMENT (FR)
BULL SCHWEIZ AKAD MED (ISS), 20:490-520, SEP 64
*ABNORMALITIES, DRUG-INDUCED, *ANTITUMOR AGENTS, *ANTITUMOR AGENTS,
*ANTIMETABOLITES, *ANTIMETABOLITES, *ANTIMETABOLITES, *ANTIMETABOLITES,
*COLCHICINE, EXPERIMENTAL LAB STUDY, *FOLIC ACID ANTAGONISTS,
*QUINETHAMIDE, RATS, *NITROGEN MUSTARD COMPOUNDS, *PARANITROPHENOL,
*PODOPHYLLIN, *PODOPHYLLIN, PREGNANCY, PREGNANCY, ANIMAL, RABBITS,
RATS, *THALIDOMIDE, TOXICOLOGIC REPORT, *TRANQUILIZING AGENTS,

MEDLARS SEARCH STATION-UCLA

Figure 9
Author Format, MEDLARS Demand Search Bibliography

ALLEGRA SR	50106	AVIOLI LV, CRACCO RQ, CHAMBERS R	50106
BRAIN SCANNING WITH MERCURY. THEORY, TECHNIQUE AND ACCURACY OF THE METHOD.		203 MERCURY BRAIN SCANS: THE USE OF SMALL DOSES AS A SCREENING METHOD.	
TURK J PEDIAT 7131-43, JAN 65		J NUCL MED 61:52-64, APR 65 **	
*BRAIN NEOPLASMS; *MERCURY; *RADIOISOTOPE SCANNING		BRAIN ABSCESS; *BRAIN DISEASES; *BRAIN NEOPLASMS; CEREBRAL ANEURYSM; *DIURETICS; MERCURIAL HEMATOMA, SUBDURAL; HYDROCEPHALUS; NEOPLASM DIAGNOSIS; *RADIOISOTOPE SCANNING; STATISTICS; SUBARACHNOID HEMORRHAGE	
BARKE R, GURSKY S	50106	HARRY D	50106
(THE RADIOIODINE TEST IN RATS) (GER)		CLINICAL ASPECTS OF BRAIN SCANNING WITH MERCURY.	
RADIOBIDL RADIDTHER (BERLIN) 5:511-5, 1964		TURK J PEDIAT 7122-30, JAN 65	
*IODINE ISOTOPES, DIAGNOSTIC; RADIOISOTOPE SCANNING; RATS; *THYROID FUNCTION TESTS		*BRAIN NEOPLASMS; GERIATRICS; *MERCURY; *RADIOISOTOPE SCANNING	

Figure 10
Card Format, MEDLARS Demand Search Bibliography

Los Angeles: Biomedical Library

The bibliography which you have received is a product of MEDLARS, a computer-based information storage and retrieval system. While the system is in its developmental phase all information concerning the operation is of value to its future success.

As a user of MEDLARS you are in a position to provide both general and specific criticism of the systems output. Since this is the area of most vital interest to you, we should greatly appreciate your comments on the attached appraisal form.

Although this form lists certain major areas which can be evaluated by checking the appropriate box, we have left space for more general comments, and we hope that you will consider giving us the fullest evaluation possible.

Again we should like to stress that you, the user, can provide the most important feedback to the system; this is what we particularly need now during the developmental stage.

Thank you for your cooperation.

MEDLARS Search Station
Biomedical Library
University of California,
Los Angeles

Figure 11
Request for Evaluation of a MEDLARS Search.

NAME _____

SEARCH NUMBER _____

DATE OF APPRAISAL _____

UNIVERSITY OF CALIFORNIA, LOS ANGELES
 BIOMEDICAL LIBRARY
 MEDLARS Search Station

Appraisal of MEDLARS Search

Quantity

Was the number of citations received greater than (), less than (), or approximately equal to () the number expected? Please note your analysis.

- _____ Search conducted on too broad a selection of subjects.
 _____ Search conducted on too narrow a selection of subjects.
 _____ Other (Please specify)

Relevance

How many citations are you aware of which this search did not retrieve?
 _____. Please specify.

What is the approximate number of citations not relevant to your search?
 _____. Is this estimate based on actual examination of the articles?
 _____.

In your opinion, what may have caused retrieval of the non-relevant citations. (e.g., descriptors used, formulation of search for retrieval)

Timeliness

Did you receive the citations in time to be of use? _____. Do you have any comments on timeliness?

General Comments (Use reverse side if necessary)

Figure 12
 Form for Appraisal of a MEDLARS Search.

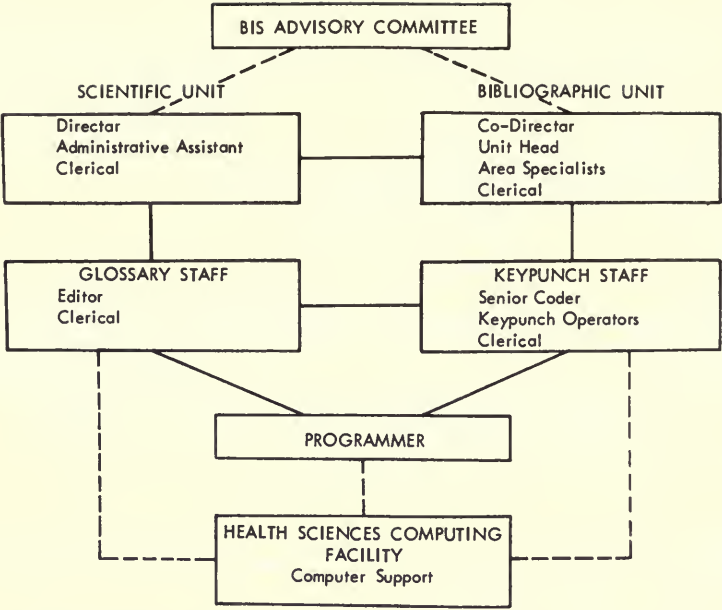


Figure 13
Organization of the Brain Information Service, UCLA

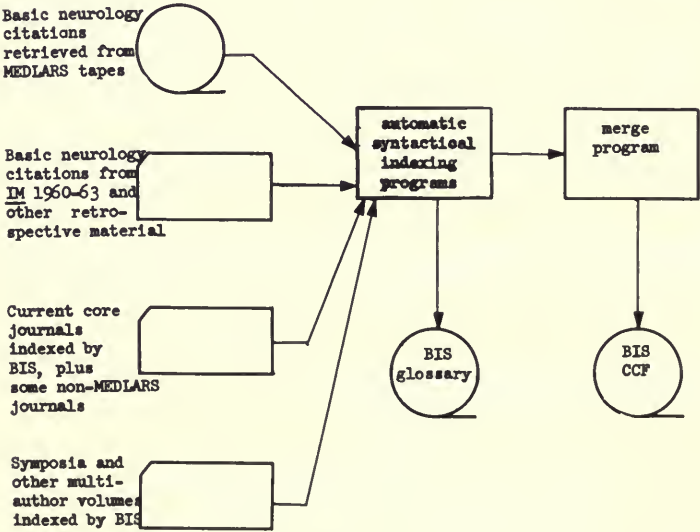
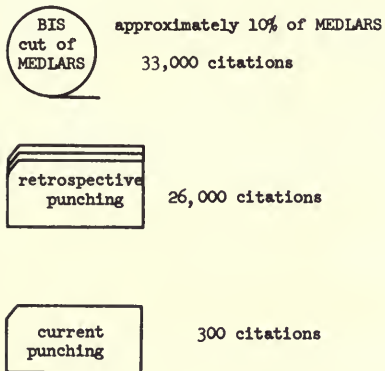


Figure 14
Input for Brain Information Service Computer Store

POTENTIAL SIZE OF BIS CCF
AS OF APRIL 1, 1966



ESTIMATED RATE OF GROWTH
PER MONTH

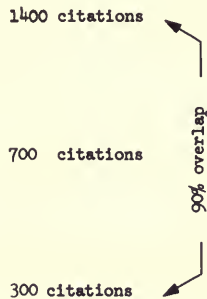


Figure 15
Size and Growth of Brain Information Service Computer Store

CITATION-ENTRY RECORD

Key punch Date APRIL 12, 1966

915-3

Indexer		1 2		3 4 5 6 7				8 9 10		11 12		Source (SEE SOURCE CODE LIST)																																																																																														
NAME		CODE		Date				Serial #		Source																																																																																																
N. NOYES		05		5 0 4 1 2				0 0 1		0 2																																																																																																
				19 66																																																																																																						
000	1	JOURNAL ABBREVIATION (37 CHARACTERS, INCLUDING MARKS AND SPACES)										03	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	0	EXP. BRAIN RES.										XXX	000	1	000	1	19	66	0	100	0	102	0	126																																																																																		
Type of article: 1. Standard, 2. Review, 3. Abstract, 4. Letter, 5. Note																																																																																																										
0	1	TERM (16 CHARACTERS)										0	1	TERM (16 CHARACTERS)										0	1	TERM (16 CHARACTERS)										0	1	FORMAT																																																																				
2	0	CORTIC-CELLICULAR										0	1	PATHWAY										0	1	ACOUSTIC C. C. R. T. E. X.										0	1	00																																																																				
2	1	AUDITORY EVOKED RESPONSES										0	1	STRYCHNINE										0	1	00																																																																																
2	2	DIRECT CNS INJECTION										0	1	NEODECORTICATION										0	1	00																																																																																
2	3	ELECTRICAL STIMULATION										0	1	POTASSIUM CHLORIDE										0	1	00																																																																																
2	4	SUBCORTICAL REFLEXO CENTER										0	1	CAT										0	1	00																																																																																
Role indication: see list. Value: Highest is 1. Format: see code list.																																																																																																										
000	1	ABSTRACT																																																																																																								
3	0																																																																																																									
3	1																																																																																																									
3	2																																																																																																									
3	3																																																																																																									
3	4																																																																																																									

Figure 16a
Citation Entry Record for Brain Information Service

	1	1	20	AUTHOR, JOUENES, JOE
13	14	15	16	50
4, 9	1			BUSER P
4, 1	2			SAINTAALAURET J
4, 2	2			ST. LAURENT J
4, 3	2			MENINI CH
4, 3				

Type of entry: 0. Link, 1. Senior Author, 2. Junior Author, 3. Editor, JOUENES, JOE, JOUENES, JOE

	1	1	20	TITLE
13	14	15	16	
5, 0	2	FR00		(INTERVENTIØN ØF THE INFERIØR CØLLICULUS IN THE ELABØRATIØN AND THE
5, 1	0	FR00		SPECIFIC CØRTICAL CØNTRØL ØF CLØNIC DISCHARGE TØ SØUND IN THE CAT UNDER
5, 2	0	FR00		CHLØRALØSE)
5, 3				
5, 4				

Type of title: 0. Link, 1. English original, 2. English translation, JOUENES, JOE, JOUENES, JOE

Figure 16b
Citation Entry Record for Brain Information Service

1314	15	16-19	20	INSTITUTION ADDRESS OR GRANTING AGENCY
80	1			(LAB. NEURØPHYSIØL. CØMPAREE, FAC. SCIS, PARIS, FRANCE)
81				
82				
83				

Type of entry: 1. INSTITUTION ADDRESS, 2. GRANTING AGENCY, 0. CONTINUATION SYMBOL

1314	15	Lang. 16-19	20	TITLE (VERNACULAR OR transliteration)
60	1	FR00		INTERVENTIØN DU CØLLICULUS INFERIEUR DANS L'ELABØRATIØN ET LE CØNTRØLE
61	0	FR00		CØRTICAL SPECIFIQUE DES DESCHARGES CLØNIQUES AU SØN CHEZ LE CHAT SØUS
62	0	FR00		CHLØRALØSE
63				
64				

Type of entry: 1. VERNACULAR, 2. transliteration, 0. CONTINUATION SYMBOL

1314	Blank 15-19	20	SCIENTIST'S NAME
90			
91			
92			
93			
94			

Figure 16c
Citation Entry Record for Brain Information Service

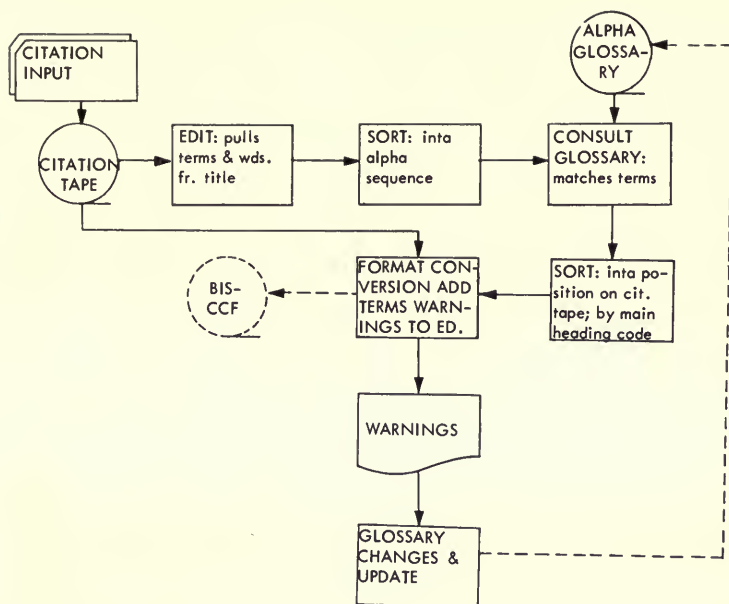


Figure 17

Flow chart showing BIS programs for input and glossary formation.

NUMERIC	BISH	MAIN	HD	CO	NR-SYN	TAG	SOURCE	MAIN	HEADING
36067	1	M						ELECTROCHEMISTRY	
36078	1	M						ELECTROCOAGULATION	
36083	1	M						ELECTROCONVULSIVE SHOCK	
36090	1	M						ELECTROCONVULSION	
36120	1	M						ELECTROENCEPHALOGRAPHY	
36120	2							ELECTROENCEPHALOGRAPHICAL	
36120	2							ELECTROENCEPHALOGRAPH	
36120	2							ELECTROENCEPHALOGRAM	
36120	2							ELECTROENCEPHALOGRAMS	
36120	2							E.e.g	
36120	2							ELECTROENCEPHALD-GRAMS	
36120	2							ELECTROENCEPHALOGRAPHIC	
36120	2							EED-STUDIES	
36120	2							ELECTROENCEPHALOGRAM	
36120	2							ELECTRO-ENCEPHALOGRAPHY	
36120	2							EES	
36120	2							ELECTRO-ENCEPHALOGRAPHIC	
36120	2							EES	
36120	2							ELECTROENCEPHALD-GRAPHIC	
36120	2							ELECTROENCEPHALOGRAPH	
36130	1	M						ELECTROGALVANISM	
36131	1	M						ELECTROGRAPHY	
36142	1							ELECTROLYSIS	
36142	2							ELECTROLYTIC	
36183	1							ELECTROLYTES	
36183	2							ELECTROLYTE	
36183	2							IONIC	
36204	1							ELECTRONYTOGRAPHY	
36204	2							ELECTRONYTOGRAM	
36204	2							ENG	
36204	2							ELECTRON SPIN RESONANCE	
36226	1	M						ELECTRON TRANSPORT	
36225	1							ELECTRONARCOSIS	
36246	1	M						ELECTRONICS	
36267	1	M						ELECTRONS	
36277	1	M						ELECTROOCULOGRAPHY	
36277	2							EDG	
36288	1	M						ELECTROPHORESIS	
36288	2							ELECTROPHYSIOLOGY	
36309	2							ELECTROPHYSIOLOGICAL	
36330	1							ELECTRORETINOGRAPHY	
36330	2							ELECTRORETINOGRAMS	
36330	2							ELECTRORETINOGRAM	
36330	2							ELECTRORETINOGRAPHICAL	
36330	2							ELECTRORETINOGRAPHIC	
36330	2							ENG	
36330	2							RETINOGRAM	
36330	2							RETINOGRAMS	
36336	1							ELECTROSHOCK	
36336	2							ELECTRO-STOCK	
36341	1	M						ELECTROSPHYNOGRAPHY	
36346	1							ELECTROSURGERY	
36351	1	M						ELECTROTHERAPY	
36362	1	M						ELEDIOISIN	

Figure 19
Print-out from the Numeric Section of BIS glossary (BISH)

INDEXED UNDER TERM FROM TERM CARD -- 098721,	RETINA
INDEXED UNDER TERM FROM TITL CARD -- 152483,	NUCLEUS
INDEXED UNDER TERM FROM TERM CARD -- 152411,	HAYWARD J
INDEXED UNDER TERM FROM TERM CARD -- 152417,	KOMISARUK BR
INDEXED UNDER TERM FROM TERM CARD -- 152430,	SAWYER CH
INDEXED UNDER TERM FROM TERM CARD -- 152450,	CLEMENTE C

THE MESSAGES BELOW REFER TO DECKLET -- 066040600602
 (ELECTRICAL RESPONSES TO VISUAL STIMULATION IN MIDBRAIN, MEDU
 LLA OBLONGATA AND SPINAL CORD OF THE LAMPREY)

THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	ELECTRICAL RESPONSES
THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	LAMPREY
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	LATENCY PERIOD
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	LIGHT STIMULATION
THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	MIDBRAIN
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	MOTOR EFFERENT PATHWAYS
THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	VISUAL
THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	VISUAL STIMULATION
INDEXED UNDER TERM FROM TITL CARD -- 069363,	MEDULLA OBLONGATA
INDEXED UNDER TERM FROM TERM CARD -- 080934,	OPTIC NERVE
INDEXED UNDER TERM FROM TITL CARD -- 106785,	SPINAL CORD
INDEXED UNDER TERM FROM TITL CARD -- 150819,	ELECTRICAL
INDEXED UNDER TERM FROM TITL CARD -- 151925,	STIMULATION
INDEXED UNDER TERM FROM TERM CARD -- 151990,	TEGMENTUM
INDEXED UNDER TERM FROM TERM CARD -- 152411,	HAYWARD J

THE MESSAGES BELOW REFER TO DECKLET -- 066040600702
 (EFFECT OF PANCREATECTOMY AND ACETYLCHOLINE ADMINISTRATION ON
 ELECTRICAL ACTIVITY OF MUSCLE AND FUNCTION OF THE NERVE-MUSC
 LE JUNCTION)

THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	ACETYLCHOLINE ADMINISTRATION
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	CONTRALATERAL INFLUENCE
THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	JUNCTION
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	LATENCY PERIOD
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	MOTOR NERVE
THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	NERVE-MUSCLE
THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	NERVE-MUSCLE JUNCTION
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	NEUROMUSCULAR APPARATUS
THIS TERM FROM A TITL CARD IS NOT IN THE BISH --	FILE AS NON-INDEXABLE -- PANCREATECTOMY
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	SYSTEMATIC ACETYLCHOLINE ADMINISTRATION
THIS TERM FROM A TERM CARD IS NOT IN THE BISH --	TROPHIC FACTOR
INDEXED UNDER TERM FROM TITL CARD -- 001134,	ACETYLCHOLINE
INDEXED UNDER TERM FROM TITL CARD -- 036309,	ELECTRICAL ACTIVITY
INDEXED UNDER TERM FROM TERM CARD -- 044709,	FROG
INDEXED UNDER TERM FROM TITL CARD -- 074256,	MUSCLE
INDEXED UNDER TERM FROM TERM CARD -- 075054,	MYONEURAL JUNCTION
INDEXED UNDER TERM FROM TERM CARD -- 150043,	ACTION POTENTIAL
INDEXED UNDER TERM FROM TITL CARD -- 150928,	FUNCTION
INDEXED UNDER TERM FROM TERM CARD -- 151818,	RESTING POTENTIAL

THE MESSAGES BELOW REFER TO DECKLET -- 066040600802

Figure 20

Print-out of Warnings and Other Messages from the System to the BIS Glossary Editor.

ELECTRICAL RECORDINGS FROM BRAIN (See also BRAIN ELECTROPHYSIOLOGY 6)

EEG WAVES, ACTIVITIES, RHYTHMS AND COMPLEXES

SLOW-WAVE EEG

DELTA ACTIVITY

DELTA RHYTHM

DELTA WAVE

THETA ACTIVITY

THETA RHYTHM (Theta-rhythm)

THETA WAVE

ALPHA ACTIVITY (See also ALPHA ATTENUATION)

ALPHA RHYTHM (Alpha-rhythm)

ALPHA WAVE

BETA ACTIVITY

BETA RHYTHM

BETA WAVE

GAMMA ACTIVITY

GAMMA RHYTHM

GAMMA WAVE

Figure 21

Table Showing a Section from the BIS Glossary Hierarchy.

REFERENCES

1. Karel, Leonard, et al. "Computerized Bibliographic Services for Biomedicine," Science, 148:766-772, May 7, 1965.
2. Darling, Louise. "MEDLARS: a regional search center." Paper presented at the 2nd Institute on Information Retrieval, University of Minnesota, November 10-13, 1965. (To be published in the Institute Papers. Much of the section on MEDLARS programs in this paper has been paraphrased for the present paper.)
3. Rogers, Frank B. "MEDLARS Operating Experience at the University of Colorado," Bulletin of the Medical Library Association, 54:1-10, Jan. 1966.
4. Garvis, Francis J. "The UCLA MEDLARS Computer System," Bulletin of the Medical Library Association, 54:14-15, Jan. 1966.
5. Planning Research Corporation. MEDLARS Users Guide, 15 November 1965. Los Angeles, 1965. (Revision in preparation. Prepared for University of California, Los Angeles, under Contract No. 70602-0.)
6. Brain Information Service, University of California, Los Angeles. "Present and projected activities; organization of the Brain Information Service bibliographic activity" [by Pat Walter]. Los Angeles, 1966. (Mimeographed.)
7. Brain Information Service, University of California, Los Angeles. "Glossary and glossary routines as of March, 1966" [by Peter Amacher]. Los Angeles, 1966. (Mimeographed.)

OPERATIONAL COMPUTER-BASED SYSTEMS AT THE CHEMICAL ABSTRACTS SERVICE

Seldon W. Terrant, Jr., and J. L. Wood

The American Chemical Society (ACS) is a non-profit organization chartered by Congress as a scientific and educational organization. The Chemical Abstracts Service (CAS) is a division of the American Chemical Society.

Like any society the ACS is concerned with helping its members to get their papers published and to help chemists and chemical engineers to gain recognition in the scientific and general public communities. Equally important, however, the ACS wants to have available information put to use. To this end, it is concerned that all forms of research, development, and applied technology be fully recorded, indexed, and retrievable for the full scientific community. Chemical information is not used only by chemists and chemical engineers. It is used by many other kinds of scientists. In line then with the major purpose of getting available information put to use, it is clear that the ACS operates the Chemical Abstracts Service not only for the benefit of members of the Society, but for a broader reason. We want chemical and chemical engineering information put to use in the full scientific community so that the general public will reap the benefits which such use will provide.

When the full dimensions of the information problem for chemistry and chemical engineering are considered, it is easy to see why a community service is a requirement for all those who use scientific information. If information processing is carried out well, the information will be cheaper and more easily available. If it is not done well, such operations will not continue to exist. Considering not only chemistry and chemical engineering information, but also all of the other scientific and technical information available, it is not possible

Seldon W. Terrant, Jr. is Managing Editor, Special Publications and Services, Chemical Abstracts Service of the American Chemical Society.

J. L. Wood is Head Librarian, Chemical Abstracts Service.

for any one institution to meet all of its information needs from within its own resources. Community services such as Chemical Abstracts are a necessity if information is to be used to the best advantage.

When the full user community for scientific and technical information and the diversified needs of the community from institution to institution and from individual to individual are considered, it can be seen that the traditional means of publication, which assemble information into traditional printed form, are neither flexible enough nor responsive enough to meet today's needs adequately. For this reason, during the last five years the ACS/CAS started to shift its operations to a computer base. It is expected that all CAS publications and services will be composed through the computer by mid-1969.

In addition to providing acquisition, analysis, organization, and indexes for all chemical and chemical engineering information, CAS intends to provide systems through which the stored data can be adequately tapped for routine use by the scientific community. The shift to a computer base will provide the additional flexibility which is necessary to be sure that the information is fully used.

Several of the CAS computer applications involve or affect the operations of the CAS Library. This paper summarizes three of these applications. First, CAS maintains a computer-based inventory of published abstracts and rejected abstract manuscripts. This is an on-line operation, designed to produce management guidance tools—tools useful not only to CAS executive management, but also tools that are very useful to the CAS Library, especially in its acquisitions work. The second is a computer-produced directory of serial and non-serial titles and their ASTM Coden. The third is a computer-based project still partially in the developmental stage. This is a computer-based list of periodicals for chemistry and chemical engineering. This is an actual scheduled project with definite internal applications and with an output that will be of direct benefit to the scientific and technical information-handling community.

Before describing our computer-based inventory system, the framework for this topic should be set by referring to a few statistics which in themselves justify the system. In 1966, CAS expects to publish approximately 225,000 abstracts of published papers and patents in Chemical Abstracts. Of these, 170,000 will be abstracts of papers and government reports which originate from over 100 countries. Approximately 50 percent of this collection is published in English, while the remaining 50 percent appears in about 50 other languages. This material will come from nearly 11,000 different serials and from between 600-700 non-serial publications, such as the proceedings volumes of conferences. It will also announce the availability of

approximately 3,000 new books. To look at this another way, assume that a reasonable estimate of the world's annual output of scientific and technical articles is one or two million articles, and that a reasonable estimate of the total number of scientific and technical journals is between 35,000 and 50,000. CAS then has the problem of abstracting somewhere between 11.3 percent and 22.5 percent of the world's annual output of scientific and technical papers and monitoring somewhere between 22 percent and 32 percent of the world's scientific and technical journals.

It is apparent then from these statistics, as well as from comparable statistics for earlier years of CAS activity, that an inventory of work accomplished is an important management tool. It has been, and still is, necessary for us to have statistical data about the number of abstracts we publish, the length of the abstracts, and the number of abstracts related to various sub-disciplines of chemistry, in order to plan for future growth and to establish staff and financial requirements. CAS needs to know the source of its abstracts for abstractor remuneration purposes. It needs to know which primary journals contained the papers that were abstracted, and the relative degree of productivity of abstractable papers of these journals. CAS needs to know how many of the abstracts are prepared in-house versus those that are prepared by voluntary abstractors, and the distribution of the abstracts throughout Chemical Abstracts.

Prior to 1963, some of these statistics were available as the result of various hand counts, while others were lacking completely. In 1963 a team consisting of persons well versed in existing CAS procedures and of systems design personnel devised a computer-based inventory system which was capable of producing the needed reports from a single input operation. Although this is an on-line operation, it is not static. It is undergoing continual change. The intention is to convert from manual to machine-based input as a part of overall mechanization of CAS operation. This discussion, however, will be devoted to the existing system.

One of the factors that enabled CAS to design the present system was the ability to represent the titles of primary publications in a highly compacted form by using the American Society for Testing and Materials Coden. CAS first adopted the use of Coden for its computer-based publication, Chemical Titles. Of the 40,000 five-character unique codes that ASTM has assigned to periodical and non-periodical titles, 24,600 of these were assigned at the request of the Chemical Abstracts Service. Coden have now been assigned to all titles listed in the Chemical Abstracts List of Periodicals back to and including the 1956 edition.

Computer-Based Inventory

The computer-based inventory begins with five documents: (1) the printed issues of Chemical Abstracts in page proof format, (2) manuscripts submitted for publication in Chemical Abstracts that have been rejected, (3) a directory of periodical titles and their unique Coden, (4) the coding sheets for printed abstracts, and (5) the coding sheets for rejected abstracts. The worksheet for printed abstracts was designed in conjunction with a set of conventions which enables clerical employees to record pertinent statistical data about printed abstracts at the rate of 700-800 abstracts per day (see Figure 1). The information across the top of the form identifies the coder, the volume of Chemical Abstracts (CA), the issue, and the section which she is coding. The form is divided into four basic parts. Part 1 is for CA reference data; Part 2, the original reference data; Part 3, the name of the abstractor or the in-house abstract code; and Part 4 is the amount of space occupied in CA by the abstract and related matter, such as structures, cross references, and blank space. Under the CA reference, space is provided to record the CA column number, the location of the abstract within the column by a letter "a" through "h," and a more specific location within the letter fraction by a numeral. The language of publication of the paper is also indicated. If the abstract exceeds a certain length, this is indicated under prior lines.

For the original reference we use the Coden to designate the journal, and add the year, volume, issue, and beginning page data. Patent numbers are recorded in the column for page data while lengthy identification such as report numbers are listed in the space for other identification. The country in which the work was done and the terminal page number are also indicated.

Under the next heading is written the last name and initials of the abstractor or a four-character code that specifies an abstract produced in-house. This four-character code indicates the source of the abstract, for example, an author abstract, a translation of an author abstract, a copy abstract, or an abstract prepared by a CAS staff member. The last column is used to note by number of lines the length of the abstract.

When coding for an issue of CA has been completed, data on the completed worksheets are transcribed directly onto magnetic tape using a Mohawk Science Data Recorder, model 1101. These data are then immediately processable in the computer to produce the following reports.

Report 1 is a density report (see Figure 2). As an editorial check, it is useful to follow closely the variation of page density of

abstracts for each of the 74 sections of CA, that is, the number of abstracts in each section and the number of pages required for each section. This variation gauges the amount of editorial effort devoted to each section. These reports are produced biweekly. They show: (1) the number of abstracts of papers and patents in each section; (2) the amount of space, in columns, devoted to abstracts of papers and patents in each section; and (3) the total number of abstracts and columns for each section. Each issue report also shows these same data cumulated for all issues of the volume being inventoried.

Report 2 is the in-house abstracting report. Each CAS abstract produced in-house is coded in CA with a descriptive four-character code. Statistics for abstracts prepared by full time CAS staff become important as new services such as Chemical-Biological Activities are offered, and as the use of author abstracts changes. This report is prepared for each volume of CA. It provides a complete inventory of in-house abstracts showing the total number in each issue (see Figure 3), and their distribution by section (see Figure 4).

Report 3, a three-part report, is designed to aid library acquisition work. It is important for us to know the sources of the papers that are abstracted and to relate these data to the acquisition and processing of the journals throughout CAS.

Report 3A (see Figure 5) lists alphabetically by Coden all of the primary and secondary journals, patents, books, and one-shot publications (conference proceedings, monographic collections of papers, etc.) represented by the abstracts in a volume of CA. Opposite each Coden is the number of abstracts from the publication. This report enables CAS to determine a cost/use ratio for the documents it purchases.

Report 3B (see Figure 6) is arranged in descending sequential order by the number of abstracts each document produced. That is, it begins with the journal which contained the most papers abstracted and ends with the one which contained the least. This report emphasizes the more productive journals and serves as a guide for establishing candidate journals for express handling. For each primary journal and each one-shot publication this report shows:

- (1) cumulative number of primary journals and one-shots,
- (2) abstracts per title,
- (3) cumulative number of abstracts of primary journals and one-shots,
- (4) percent of total number of abstracts at 5-6 percent intervals,
- (5) cumulative number of primary journals,
- (6) cumulative number of abstracts of primary journals,
- (7) percent of total number of abstracts for primary journal papers based on 5-6 percent intervals,
- (8) cumulative number of one-shots,

- (9) cumulative number of abstracts of one-shots, and
(10) percent of total number of abstracts for one-shots based on 5-6 percent intervals.

The tapes used to generate these volume reports are merged, and listings of journals with cumulative productivity data are printed. These cumulative lists give us a truer picture of the journals which are most productive, and these are the journals we handle on an express basis. By cumulating data from these reports for four years, and plotting the number of journals required to produce any given percentage of the total number of abstracts, an interesting curve is obtained (see Figure 7). The curve shows that 25 percent of the total number of abstracts were obtained from 50 journals; 50 percent were obtained from 250 journals; 75 percent were obtained from 1,000 journals; while 90 percent came from slightly over 3,000 journals. The remaining 10 percent of the abstracts came from an additional 2,500 journals. We monitor an additional 8,000 journals of which an average of 2,500 produce the final 10 percent of the total number of abstracts.

Report 3C (see Figure 8). The growth in many areas of chemistry and chemical engineering is gauged by monitoring other secondary journals that overlap CA coverage. We currently monitor approximately 50 such secondary publications. This report lists the primary journals that contain papers covered as copy abstracts.

Report 4 is used to generate the CA abstractors' payroll. The output is the abstractors' actual pay checks. CAS pays its abstractors by the line as printed in CA. Therefore, it is necessary to determine for each abstract the number of lines and to record this information along with the abstractor's name. The payroll is produced on a quarterly basis.

Report 5. Certain areas of chemistry produce a great many index entries per abstract. For instance, papers dealing with chemical synthesis report a great many compounds. The corresponding sections of CA are indexed directly from the original journal article. Thus for the indexers of these sections their primary document needs can be predetermined. One of the products of the inventory is the direct print-out of library call slips (see Figure 9) for those sections which have a high density of index entries per paper. This permits the library to arrange the original journal material in a manner convenient for the indexing effort.

Report 6 is a listing of document titles, with the number of abstracts from each, by subject based on CA's 74 sections. Either single sections or a combination of sections can be listed. Coverage within each of the individual sections of CA is continually watched. In addition, each year we select several sections to investigate in depth. Here we review with the section editor those subdivisions

which he feels deserve more attention, and the primary journal sources for each subdivision are investigated.

Report 7 is a listing of references to printed abstracts arranged according to bibliographical citations by primary journals. It is used to check coverage of a given journal. When merged with Report 8, which is described below, it serves as a useful tool to search for holes in the coverage. We intend to extend the usefulness of this report by inputting data about papers that have been perused but not selected; thus, any gaps in the coverage will be very apparent; and, in fact, the report will take a negative approach to the problem, i.e., it shows us what has not been entered into the CAS system.

Report 8. Another aspect of the computer-based inventory is the record of abstract manuscripts that have been rejected. Each year a number of the abstracts submitted to us for publication do not meet the criteria for publication in CA. For example, they do not report chemical information or new chemical information. Others are duplicates of abstracts that have already been published. Duplicates fall into two classes: those papers which are abstracted twice by error and those papers that are published in two or more sources. These rejected abstracts are coded on worksheets (see Figure 10), transcribed onto magnetic tape, and reports are generated which show (1) the primary journal address of the article, (2) the name of the abstractor, (3) why the paper was abstracted, that is, did the abstractor do it voluntarily or did we ask him to do it, and (4) why the abstract was not used.

CAS Coden Directory

A basic tool for all of our computer-based work which involves the titles of journals and abstracted monographs is the CAS Coden Directory. This Directory now contains an alphabetical listing of serials and nonserials by full title, each accompanied by a unique five-character Coden with a numerical sequencing or alphabetizing number (see Figure 11). The Directory has an index arranged alphabetically by Coden. The data contained in the Directory is all in upper-case type. We are now in the process of writing programs which will enable us to revise the Directory. The new Coden Directory will contain (1) the Coden, and in upper and lower case type (2) the full title, and (3) the American Standards Association title abbreviation (see Figure 12). This Directory will also contain a numerical sequencing number and will be accompanied by an index arranged alphabetically by Coden. This Coden Directory constitutes an important source of input for the periodicals list as will be explained later on.

Periodicals List

Before describing the work that we are doing on our new List of Periodicals, the history of this continuing project is worthy of mention. CAS has always paid conscientious attention to nomenclature. This applies not only to the names of chemical compounds, but also to the titles of chemical publications. In 1906 the American Chemical Society published a list of journal titles in the Proceedings of the American Chemical Society. It was entitled "A List of Abbreviations" and contained 37 journal titles and the abbreviated form for each. This list was revised, enlarged, and issued as a separate publication under the promising title List of Journals to be Abstracted by Chemical Abstracts. It preceded by several months the first issue of Chemical Abstracts and consisted of abbreviated titles for 370 journals. In 1908 the list was again revised and published as a part of the introduction to the second volume of Chemical Abstracts. It contained 475 abbreviated titles and was accompanied by three simple rules pertaining to journal title abbreviation. In 1910 the list appeared as the List of Periodicals, and the form of entry still in use today was introduced. The full title was given with those letters constituting the abbreviation set in bold face type. In addition to the title and abbreviation, the entries included frequency, volume number, price, and the publisher's address. The List continued to be issued biennially until 1922, when a second important feature was added—library source-guide data for 162 American libraries. This key to library files gave new value to the List, for it now became a useful tool for locating abstracted journals in the libraries of the United States.

The purpose of the List was two-fold. First, it was to standardize the abbreviated title form. Second, it was to furnish bibliographic data about abstracted papers by providing a library source guide. Beginning in 1936 the List was published quinquennially. In 1957 the first annual supplement to the quinquennial edition appeared.

CAS has now begun work on a completely new edition of the List of Periodicals. This new List is being designed to include much more data than did earlier editions. It is to be made available in both printed and machine-readable form (either punched cards or magnetic tape). Programs for searching the tapes will also be made available.

This catalog of periodicals related to chemistry and chemical engineering will bring into a single source much valuable bibliographic data in considerable detail that is now only partially available and from a large number of bibliographies, union lists, and individual library files. A typical entry in the List will include the full title of the publication; a translation of the title if the language is other than English, French, German, or Spanish; the ASA Z39 standard title abbreviation; the ASTM Coden; the languages of publication and

summaries; the history of the periodical with references to any former titles, current frequency, and volume number data; the publisher's address; and the price. The entries will also include the title cataloged according to American Library Association cataloging rules and the library holdings information (note that we will list holdings rather than merely an indication that a library has a current subscription to a journal).

Initially we intended to publish a 1966 edition of the Chemical Abstracts List of Periodicals with Key to Library Files. This List would have been merely an updated version of the 1961 List. The relationship between it and the 1961 List would have been identical to the relationship between the 1961 List and the 1956 List, that is, already published information about discontinued periodicals and name changes would not have carried forward into the new List. Had this course been followed, the 1966 List would have been the 14th CA List of Periodicals. To do a literature search using all of Chemical Abstracts, one would have had to use all 14 editions. As an alternative we investigated the possibility of publishing a 60-year cumulative edition of the List of Periodicals, thus bringing within one set of covers pertinent bibliographic data about, and library source guide information for, the serials listed in all previous editions.

The next logical extension of the List was to consider the inclusion of data for other serial publications related to chemistry and chemical engineering which had not been abstracted by Chemical Abstracts. A survey of Beilstein's Handbuch der Organischen Chemie revealed some 300 pre-1907 titles. Coverage of the literature of pure and theoretical chemistry from 1830 through 1940 by Chemisches Zentralblatt (CZ) was particularly outstanding. We discovered approximately 500 journals covered by CZ prior to 1940 had not been covered by CA. We also developed a list of approximately one hundred defunct nineteenth-century chemical journals to be included.

Thus the List which is proposed will include approximately 24,000 entries for journals, and 5,000 entries for monographs abstracted by Chemical Abstracts, Chemisches Zentralblatt, and covered by Beilstein, as well as the nineteenth-century chemical journals never covered by any of these services. This extension makes our List truly a comprehensive list of periodicals for chemistry and chemical engineering.

The procedure used to generate this List may actually be viewed as two separate projects: (1) gathering the pertinent bibliographic data, and (2) obtaining library holdings information from some 350-400 cooperating libraries. Our first step was to establish a file of the serial and nonserial titles. This was done by cutting and pasting the entries from all of the Lists of Periodicals, entering them into a single file, and adding to this file titles for periodicals from Chemisches Zentralblatt, Beilstein, and several other sources. Once the

file was established, the Coden was added to each main entry. This necessitated the acquisition of several hundred additional Coden from ASTM (see Figure 13).

The first publication our work generates is a checking edition for the List of Periodicals (see Figure 14). The checking edition contains the titles cataloged according to ALA rules, the Coden, the history of the serials, and a space for the cooperating libraries to write in the dates of their consecutive holdings. It is necessary to furnish the cooperating libraries with a checking edition arranged according to ALA cataloging rules, since the majority catalog their serials according to these rules. The checking edition for the CA 1961 List was produced on unit card equipment. It contained 9,000 entries in ALA form. Fortunately, these cards had been retained.

A new checking edition is now being prepared in three parts for use by the cooperating libraries. The punched card with the ALA form of entry is matched with 5" x 8" cards from the master file. If no punched cards are available, the ALA form of entry is written on the 5" x 8" card. Historical data are added and the cards passed to keypunchers to generate a new set of punched cards containing the Coden, the title in ALA form, and the history. These cards are inter-filed. When the file is complete, they will be converted to tape. The resulting print-out will be used to produce camera-ready copy for printing the checking editions.

These checking editions will be sent to the cooperating libraries, and when each is returned, cards will be punched showing the Coden, the cooperating library's code number, and the library's consecutive holdings of the serial (see Figure 15). These cards will be read onto tape, sorted, and a tape containing library holdings information will be ready for future use.

There will be then one tape with the full title of a publication in upper and lower case, its official ASA Z39 abbreviation, and ASTM Coden. On another tape there will be the ASTM Coden, the ALA form of entry for the publication, and the history data.

The cards in the master record file, which contain Coden and entries for the publications as they appeared in a CA List of Periodicals, will be compared to current issues of the serials and nonserials. The entries will be edited, expanded, and updated (see Figure 16). The data on the edited master record file card will then be keyboarded on a Dura Mach - II paper tape generating typewriter for later conversion to magnetic tape. The worksheet is being designed so that the Dura typist will not have to type any of the information already in machine-readable form, i.e., she will not have to keyboard the title or abbreviated title, the history, or the ALA form of entry. These will all be retrievable by inputting the Coden. She must add needed subtitles, English translations, references to the former titles, languages of publication and summaries, the frequencies of

publication, the current volume numbers and years, prices, and the identification of the publishers. Repetitive phrases such as "Printed in," "with_____ summaries," etc., will be stored in the computer so that these do not have to be keyboarded but only coded.

Entries for discontinued periodicals, for former titles of periodicals, and for congress and symposia proceedings volumes will contain those elements of information that are pertinent to them. For instance, entries for defunct titles will not contain publisher or price data. For congress and symposia proceedings volumes, the number of the meeting, if it is one of a series, the place where it was held, the date of the meeting, the number of volumes of papers published, and the editor's name will be entered.

After the paper tape has been generated and converted to magnetic tape, this will be entered into the computer, matched with the data from the Coden directory tape, data from the tape containing the ALA form of entry, and the history information. The output will be a proof copy which will then be checked against the original edited 5" x 8" card. If any corrections need to be made, the proof copy will be edited and a corrected paper tape generated and recycled. The result of this work will be the bibliographic data tape. This tape will be merged with the library holdings tape and a master record tape generated from which a print-out will be produced for camera-ready copy for the printed version of the periodicals list.

From this master record tape, it will be possible to generate special listings of the journals by specified characteristics such as country of origin, language, price, etc., or by any combination of the elements which constitute the complete entry. Library holdings and union lists will also be obtainable (see Figure 17). The present production schedule calls for the completion of the Comprehensive List of Periodicals by the end of October 1968.

Chemical Titles

The CAS computer-based system really began in 1961 with the publication of Chemical Titles (CT), the first production-scale application in the world of computers to produce an index to technical papers. CT was the forerunner of a whole new generation of mechanized services. CT is a biweekly, computer-based publication which covers titles of chemical and chemical engineering interest taken from 691 journals. Approximately 90,000 titles were published in CT in 1965. All articles covered in CT are also covered in CA. The majority of the journals covered is received at CAS in advance form and thus the coverage is quite prompt. The titles are selected by the professional staff at CAS, translated, and then edited. Editing

of the titles includes segmentation of words to facilitate scanning of the indexes by users. The article titles, Coden (a five-letter code for the journal title) references, and authors are keypunched into cards, and the information then transferred to magnetic tape for computer processing. Check lists (corresponding to galley proofs in standard publication procedures) are printed by the computer, and the lists are proofed to insure the accuracy of the keypunched data.

Each two weeks the input data is processed by the computer and sorted into the three parts of CT: the Keyword-in-Context (KWIC) Index, the Bibliography Section (tables of contents), and the Author Index for the issue. The information is printed by the computer on graphic arts quality paper and is composed into pages mounted on special cardboard backing. The bold face journal titles are prepared on a Varityper Headliner and pasted on the pages. The camera-ready copy is then sent to the printer, where it is photographically reduced and printed on an offset press.

The KWIC index is an alphabetical arrangement of keywords selected from the article titles (see Figure 18). The keyword appears in the center of the page with its associated words preceding and following it in the same manner as they do in the title itself. Since the number of words surrounding the indexed keyword is limited by the breadth of the page, words or phrases are sometimes truncated. If the keyword appears near the beginning or ending of a title, the title will "wrap around" to fill the line space. A "wrap around" is indicated by a small cross in the text. Each keyword index entry carries a reference to the Coden, volume (or issue), and page of the journal in which the title can be found (and by which entries are listed in the bibliography section). The KWIC index provides an average of six entries to each title. The bibliography section lists the journals arranged alphabetically by Coden (see Figure 19). For each Coden, the journal title, volume, and issue are given. The titles taken from each journal are grouped together. The author index lists all authors alphabetically (see Figure 20). Each index entry provides a reference to the bibliography section.

CT is thus arranged to permit the location of information of interest via keywords, authors, or journals. It is apparent that the magnetic tapes from which the issues of CT are produced can be useful in special subject searches. During 1964, several cooperative experiments were conducted involving CT tapes. In one of the experiments, individual scientists at Eli Lilly and Company provided search terms covering their individual interests. Copies of the magnetic tapes, corresponding to each CT issue, were delivered to the company and computer searches were performed. Search results in the form of individualized listings of papers were given to the participating scientists. Each individual was allowed to change his set of search terms as his interest became better defined or as it changed. In

another experiment, searches were performed at CAS for the Olin-Mathieson Corporation. In this program the set of search terms represented the interests of a research group.

The highly successful results of these experiments and the substantial interest shown by other organizations provided the basis for the introduction in 1965 of subscriptions to CT tapes and searches. A subscriber may receive copies of the magnetic tapes or he may have his searches run by CAS in Columbus, Ohio, with the search results supplied in the form of a computer print-out. The tapes and searches are provided at the time that CT copy is sent to the printer—about a week ahead of the appearance of the CT issues. Along with the tapes are provided search programs written for an IBM 1401 or 1410. These programs are now being rewritten for the IBM System 360 computers and are expected to be available by May, 1966. The program can handle “and,” “or,” and “not” logic. It is also possible to specify word fragments as search terms. These aspects of the program will be explained in detail later. The search product can be any or all of the following: a bibliography of the selected titles, a KWIC index of the selected titles, or an author index based on the selected titles.

As part of the production of CT, a Coden-volume checking feature is incorporated. The check is based upon sets of valid Coden and Coden-volume relationships stored in the computer. Each time that information is input to the computer, it is tied to the appropriate bibliographic data. A computer check is made to determine the validity of these data and a list of invalid relationships is printed with each check list. In this manner errors can be detected prior to the production of camera-ready copy.

Chemical-Biological Activities

It has been said and written many times in recent years that the sheer bulk of primary literature being published makes it almost impossible to keep abreast of new information in a given field. This is particularly true in the biochemical fields. The number of abstracts published in the biochemical sections of Chemical Abstracts in 1965 is estimated at 63,000. This is 32 percent of the approximately 195,000 abstracts published for all fields. The figure is based on coverage of approximately 11,000 journals, in 50 languages, from 100 countries.

When the magnitude of this amount of information is considered, it becomes readily apparent that specialized alerting and retrieval services are needed. To help meet this need, CAS introduced Chemical-Biological Activities (CBAC) in January 1965 to provide a

rapid survey of literature dealing with the biological activity of organic chemical compounds. CBAC is a biweekly, computer-based publication issued in two volumes annually.

In the development of CBAC, a number of features were recognized as being desirable for a current-awareness publication in the field of chemical-biological activity. Among these are: promptness; adequate coverage of pertinent journals; information on the journal sources, including names and addresses of the authors; a format that combines legibility with ease of scanning; brief but adequate digests of the work reported (including descriptions of the substances employed, structural formulas, conditions under which the observations were made, and an account of the results obtained); and finally a set of indexes and a method of storing data such that the indexes can be cumulated.

The developmental work on CBAC, including the original concepts, the systems development, and the necessary computer programming, was done by the CAS Research and Development Division, and this division continues to assist as the system is modified and improved.

CBAC is a service which combines the values of an alerting tool, through printed abstracts and indexes of new developments in biochemistry, with computer-based storage and retrieval of information. As such, CBAC's function is broader than that of either a repository for information or an abstract publication alone. In the latter capacity, as has been mentioned, it is a biweekly publication. Its abstracts are known as "digests" to distinguish them from those published in Chemical Abstracts. Since a considerable portion of incoming material earmarked for CBAC is received by the Library in advance form, the coverage is quite prompt.

Considered on a broader scale, CBAC is not only a self-sustaining entity but is an integral part of a larger system for specific chemical information retrieval, the CAS Chemical Compound Registry. The Registry is based on the assignment of individual numbers (Registry Numbers) to each organic compound of known structure which enters the system. The Registry Number assigned to a compound of known structure identifies it completely—different ionic forms, isomers, and compounds containing unusual isotopic masses are given unique numbers. Use of the numbers provides a specific method of storing and retrieving information on any registered compound within the CAS system. CBAC, then, is a tool whose purpose is to make the literature of its defined fields more readily available to the scientist, both immediately and subsequently.

CBAC serves a field in which there is a great deal of industrial, academic, and governmental interest. At present, digests are prepared from articles taken from 350 English journals, 70 German, 38 Russian, 35 French, 18 Italian, 17 Japanese and 20 journals in 7

other languages. These primary journals are examined as rapidly as they are received at Chemical Abstracts Service.

To be defined as suitable for coverage in CBAC, a scientific paper must describe original work in one or more of the following areas: first, the effect (or effects) of organic compounds on the physiology or biochemistry of systems in or derived from plants, animals, or micro-organisms; second, metabolism of endogenous or exogenous organic compounds by the organisms just enumerated, or by systems derived therefrom, and third, in vitro chemical reactions, in the absence of a biological system, between organic compounds of biochemical interest.

A more detailed discussion of the criteria used in selecting articles for CBAC is appropriate. The first category, the effect of exogenous organic compounds on organisms or systems, probably applies to the largest number of papers selected for CBAC. This covers portions of the work done in almost every field of biochemistry, for example, enzymology, endocrinology, protection against irradiation, microbial biochemistry, mammalian and non-mammalian biochemistry, plant pathology, immunochemistry toxicology, pesticide and herbicide studies, and pharmacology.

The second category, dealing with metabolism, includes work in much the same areas mentioned, with the difference that here the biological organism or system is exerting an effect on a compound rather than vice versa. Since this action is usually described or characterized by alterations in chemical structure, this type of work is generally more chemical and less physiological in nature than the first category.

The third category, concerning in vitro reactions between organic compounds of biochemical interest, deals with those compounds not directly originating from a biological system in the particular experiment. It includes, for example, studies carried out using commercial enzymes in synthetic systems, as contrasted with the use of directly biological enzyme systems contained in homogenates or metachondria. Examples of other types included in this category are chemical treatment of proteins to elucidate their structures (when some biological activity is implied) or rupture of nucleotide bonds by chemical treatment in vitro, which is of interest in the mechanisms of mutations.

It need scarcely be pointed out that the interest of a given paper may overlap two or more of the three categories. For example, work reporting the effects of an antigen on a host, plus the fate of the antigen, would overlap two categories. This has no bearing on the way in which information is abstracted; all chemical-biological information is treated in context. As a result of the biological nature of the studies, the vast majority of the compounds encountered are organic. However, when pertinent, elements or inorganic compounds are dealt

with in the text. For example, excretion of potassium or other inorganic substances under the influence of a hormone or diuretic would be covered.

CBAC is divided into four sections, viz., the digest section and three indexes—the keyword-in-context (or KWIC) index, the molecular formula index, and the author index. The digest section contains the core information (see Figure 21). The digests are printed in such a way as to permit rapid scanning for subjects of interest to the reader. The names of organic compounds or classes of compounds appear at the left margin of the text lines so that their names may be read by moving the eye down the page. To the immediate left of the name for each fully defined compound, its unique Registry Number appears. The Registry Number appears only once for each compound within a given digest.

In addition, each digest contains some words which are highlighted by being printed entirely in capital letters. These are terms which, in addition to the organic names, are felt to be of interest in rapid scanning. They include, for example, names of elements and inorganic compounds (as these do not necessarily appear at the left margin), diseases and other pathological conditions, names of microorganisms (including viruses), names of organs or other anatomical compounds (mitochondria, liver), and words which indicate a function of interest in the digest (for example, oxidation or diuretic effect).

As an additional aid in scanning, structural formulas are given in the digests. These are selected by the chemists who edit the digests. The selection is based on providing a clue to the important type (or types) of organic compounds dealt with in the digests. For example, if a number of phenol analogs are discussed in a digest, only one example of a phenolic structure would be selected for illustration. A "large" structural formula is given only once in an issue and subsequent occurrences of the compound represented are cross-referred to the first appearance within the issue.

All papers selected from a given journal are printed under a bold face group heading which gives the journal name, volume, issue number, and year. Journals are listed alphabetically by Coden, rather than by full title. Each digest is headed by the title of the original paper (if in a language other than English, the title is translated). Following the title appears the complete bibliographic reference, including Coden notation, volume number, issue number, year, and inclusive page numbers. Names of authors are then given, by last name and initials, in the order given in the journal. Finally, the research site is given, when known.

All digests within a volume of CBAC (13 issues) are numbered consecutively. Each sentence in the digest is separated from the preceding sentence by a blank space. All sentences except the first are given subnumbers. The numbers and subnumbers are used in the

KWIC and molecular formula indexes to give a precise reference to that portion of the digest where the indexed entry appears. Keywords are selected by the chemist as he prepares the digest. As its full name implies, the KWIC index (see Figure 22) presents keywords in the context of the sentence within which they appear. Thus, the user is not limited to the information contained in the single indexed word ("actinomycin," for example, may be indexed several times) but he can, within the limits of a phrase, obtain an idea of the context in which the word is being used. The indexed entity appears in all-capital type in the center of the page, with its associated words preceding and following it in the same manner as they do in the digest itself. All index entries refer the user to the digest number or sub-number from which they were generated. The KWIC index provides an average of 30 entries to each digest.

The molecular formula index presents an ordered listing of the molecular composition of organic compounds to which Registry Numbers have been assigned (see Figure 23). To the right of the formula appears the compound's Registry Number; to the left the number (or numbers) of the digest (or digests) in which the compound is mentioned. The reference is given to the digest number (or subnumbers) in which the compound first appears; subsequent listings of the compound within the same digest are not indexed. The molecular formula and Registry Number do not appear opposite each digest reference, but only opposite the first reference; subsequent references are opposite a blank space. Molecular formulas are listed in order of increasing numbers of carbon atoms, and each formula is arranged according to a modified Hill system—carbon first, hydrogen second, and then other elements alphabetically by atomic symbols. (Note that in this system, CCl_4 , which has no H atoms, will be listed before CHCl_3 or pure hydrocarbons such as C_2H_4 .) However, the molecular formulas of metallic salts, amine salts, and addition compounds are listed immediately following the molecular formulas of the parent substances. This is done in order to eliminate the separation that would otherwise occur between parent and derivative. Thus, following the molecular formula of a carboxylic acid, references to its sodium salt would appear. Even though no actual reference to the parent compound may appear in the given issue, its molecular formula will be given in the index to facilitate finding the molecular formulas of derivatives which are referenced.

Finally, the author index presents an alphabetical listing of all author names in the issue, each name being matched with the corresponding digest number (or numbers). The three indexes are cumulated and issued every six months (corresponding to volume indexes).

The steps involved in producing CBAC are briefly as follows (see Figure 24). Journals which are covered in CBAC are scanned in the Assignment Department. A biochemist reviews each article in

depth to determine if it deals with one or more of the subject criteria used in the selection of CBAC oriented material. Digests of the articles selected are prepared by chemists at CAS. All articles covered by CBAC are also covered in CA. The digests, which are on standard forms, are sent to our Formula Indexing Department where registry sheets (containing structural formulas) are prepared for each of the organic compounds with defined structures. The digest information is then keypunched into cards and entered on magnetic tape by means of the computer system. Paralleling the processing of digests, the registry sheet information is being processed in the CAS Registry System, where Registry Numbers are assigned to the compounds by the computer. These numbers are stored on tape for merging with the corresponding compounds in the text of the digest. Computer worksheets for each of the digests are edited and any corrections re-entered onto the computer tape. The information on tape pertaining to the digest section is rearranged by the computer to provide the KWIC index, the molecular formula index, and the author index. Pages of the digest section and the indexes printed by the computer are pasted on cardboard sheets and photographed. The film is used to prepare plates for offset printing of the CBAC issues.

CBAC is constantly being evaluated for possible improvements. A case in point is an additional index included in the cumulated indexes of Volume 2 (end of 1965). This is a Registry Number and Registry Number Cross-Reference Index that will not only permit a search of the digests for individual compounds, but also for compounds whose structural features are similar. For example, an amine will be cross-referred to its acid salts by Registry Numbers and vice versa.

Starting in January of 1966, CAS began to provide, on a subscription basis, the magnetic tape records of CBAC, to be searched at the subscriber's location or at CAS. The tapes include all of the information contained in the digest section, as well as the molecular formulas of the registered compounds. This permits searching of the tapes for chemical compounds by names and by Registry Numbers and also searching for syllables, suffixes, prefixes, words, or phrases according to a given interest profile. Thus, it will be possible to conduct a search for chemical-biological activity correlations. The output of the search will be bibliographic references (CBAC digest number, original journal by Coden, volume, issue, and page, as well as titles and authors.) A KWIC index of the titles cited or a list of the authors will also be available, if desired.

One of the features that will permit great flexibility in searching for compounds is the ability to use Registry Numbers as search terms. This method will permit location of registered compounds in digests, regardless of the chemical names used to describe the compounds. The search programs permit "and," "or," and "not" logic.

Weighting factors are available in the search programs to increase the relevancy of the data recovered.

The present search programs are designed for use with an IBM 1401 or compatible computer. However, the search is being reprogrammed in a higher language, and it is expected that it will be complete by May 1966. The new programs will be oriented toward the IBM system 360 family of computers and should permit reprogramming to utilize other computer configurations. The same Coden and Coden-volume computer checks as discussed in conjunction with CT are applied during the production of CBAC.

CAS Chemical Compound Registry

It was mentioned previously, in the discussion of CBAC, that Registry Numbers provide a specific method of storing and retrieving information on registered compounds within the CAS system. The Chemical Compound Registry is the heart of the CAS computer-based system. The Registry files must contain a detailed description of each compound. At the minimum, the computer record of each compound includes the full detail of the conventional structural diagram. Chemical nomenclature does not form an adequate basis for the computer record. The Registry file also supplies the basis for substructure searches—searches for fragments of molecules. Again, nomenclature does not provide enough details for direct substructure searching in depth. It is important that the computer files include no undetected synonymy, which if uncontrolled would increase the files to unmanageable proportions.

A tabular description of the structural diagram of compounds is used in the Registry. The description, which is called a connection table, lists each of the atoms in the molecule atom-by-atom and bond-by-bond. Programs have been developed to manipulate this tabular description so as to convert the table to a unique, unambiguous form. The form is not quickly intelligible to the chemist, but it is easily handled by the computing system. Thus, the Registry file of structures consists of an ordered list of unique, unambiguous descriptions of the individual compounds. The registration process amounts to an ordering and merging by computer. The purpose is to determine whether a particular structure has already been stored in the system. If a compound has been registered previously, the same Registry Number is supplied as output; but if a compound has not been registered, a new Registry Number is provided.

Future Computer-Based Operations

CAS has plans in preliminary design state for several automated library systems capable of handling a large input of documents on a very short time-span basis. This effort is but a part of the overall work being done to put CAS on a computer base.

The master record tape produced for the publication of the Comprehensive List of Periodicals will serve as the basis for automating many of our existing library systems (see Figure 25). Existing serials inventory files, source files, disposition data files, and coverage data files are to be converted into machine-readable form and these data are to be organized in disk storage. Once the data are in machine-readable form, it will be possible to make many uses of them. The library applications include automated serials check-in and rapid retrieval of routing or disposition information. At CAS the latter is extremely important. Misrouted serials can result in duplication of abstracting or in papers not being abstracted at all. It now takes between six and nine months to train fully a clerical employee to record and route incoming new materials with minimum supervision. Under our new system this training period can be reduced by the length of time it takes to train the clerks to inquire of the computer what it is that they are supposed to do with the issues at hand.

We will be able to provide printed catalogs of the serial holdings and receipts with coverage data. We also intend to reproduce our existing dictionary card catalog in machine-readable form and to produce printed book catalogs. This process will also generate the book announcements for publication in Chemical Abstracts.

SEC	VOLUME 84 STATISTICAL REPORT				ISSUES 1 THRU 4				TOTALS			
	PAPERS		ISSUE NUMBER 4		PAPERS		ISSUE NUMBER 4		PAPERS		ISSUE NUMBER 4	
	ABST	COLUMNS	ABST	COLUMNS	ABST	COLUMNS	ABST	COLUMNS	ABST	COLUMNS	ABST	COLUMNS
1	36	2,57	00	36	2,57	141	10,74	99,88	00	141	10,74	10,74
2	134	22,53	3	207	24,90	659	49,88	33	3,89	692	492	103,77
3	207	32,53	3	207	32,53	246	49,88	1	1	647	99,83	99,83
4	76	14,58	4	53	15,11	246	49,88	1	1	12	12	12
5	90	10,12	6	1,20	56	11,32	146	28,12	19	5,09	125	38,65
6	54	16,64	00	94	16,64	376	60,91	00	00	376	60,91	60,91
7	40	6,34	00	40	6,34	155	23,09	00	00	155	23,09	23,09
8	180	29,28	4	7,47	184	29,75	507	75,39	14	2,38	521	77,77
9	308	44,56	43	7,21	308	50,47	917	146,45	113	19,96	1030	166,41
10	308	44,56	43	7,21	308	50,47	917	146,45	113	19,96	1030	166,41
11	40	6,12	33	6,45	73	12,57	120	182,90	131	32,17	160	12,42
12	131	22,58	00	131	22,58	1003	143,13	00	00	160	12,42	12,42
13	125	20,56	18	3,15	143	23,71	603	86,94	70	11,74	673	98,68
14	129	21,35	00	129	21,35	416	70,47	00	00	416	70,47	70,47
15	140	22,05	50	7,44	190	29,49	708	125,49	250	39,91	958	165,40
2000	337,47*	2222	26,94*	2222	364,41*	7826	1,216,93*	661	118,64*	6487	1,335,57*	1,335,57*
16	186	23,73	86	11,75	272	35,48	607	80,19	170	22,12	777	102,31
17	29	5,23	76	11,96	105	15,48	116	18,88	238	37,96	354	56,84
18	58	9,92	23	3,78	81	13,70	189	30,63	69	10,09	258	40,72
19	150	27,44	46	8,72	196	36,16	590	107,68	143	24,64	733	132,32
20	134	13,18	42	1,05	206	34,23	544	89,67	147	25,67	691	115,34
21	104	13,18	42	1,05	206	34,23	544	89,67	147	25,67	691	115,34
22	184	2,04	15	1,97	133	4,01	366	18,31	119	20,45	485	78,62
23	40	6,02	9	1,33	49	7,35	150	24,18	28	3,85	178	23,05
24	44	5,25	7	0,87	51	6,10	208	28,55	35	4,88	243	33,43
25	129	19,94	00	129	19,94	774	130,33	00	00	774	130,33	130,33
26	44	7,02	8	12,77	52	7,79	122	20,25	23	2,92	145	23,17
27	18	14,25	68	12,77	186	27,11	279	47,98	270	50,99	549	98,97
28	28	2,08	8	1,39	25	3,47	56	20,66	36	7,28	134	27,94
29	17	2,08	8	1,39	25	3,47	56	20,66	36	7,28	134	27,94
30	87	13,44	34	4,52	121	17,96	247	34,18	116	17,90	363	50,05
1161	184,92*	489	79,62*	1650	264,54*	4415	712,46*	1441	235,61*	5856	948,07*	948,07*
31	6	39,20	00	6	39,20	23	1,70	1,70	4	1,46	27	3,16
32	180	14,22	105	18,96	182	33,12	680	142,16	439	80,46	680	142,16
33	57	14,22	105	18,96	182	33,12	680	142,16	439	80,46	680	142,16
34	19	10,58	17	4,62	36	15,20	87	29,33	89	91,89	178	130,06
35	63	17,57	84	25,36	147	42,93	305	62,93	350	91,89	655	178,95
36	14	4,63	17	6,55	31	11,18	84	22,90	74	20,14	158	43,04
37	64	17,95	35	11,86	99	29,81	266	90,35	265	85,25	531	175,60
38	110	41,21	70	26,41	180	67,62	285	102,13	287	101,60	572	203,73
39	29	10,53	21	4,13	61	14,56	177	41,04	103	21,78	280	62,82
40	29	10,53	21	4,13	61	14,56	177	41,04	103	21,78	280	62,82
41	12	6,61	5	3,43	17	10,04	194	28,38	13	3,80	121	34,18
42	27	7,71	39	17,60	66	25,31	82	26,37	95	41,08	177	67,44
43	33	9,30	8	1,60	41	10,90	161	45,02	39	7,22	200	52,24
44	35	10,47	8	3,30	43	13,77	112	33,84	37	14,05	149	47,89
689	200,93*	410	123,99*	1099	324,92*	2654	731,66*	1809	495,56*	4463	1,227,22*	1,227,22*

Figure 2
Example of Density Report of Chemical Abstracts

IN HOUSE ABSTRACTORS BY CODE FOR VOL 63

02/21/66 PAGE 01

CODE	NUMBER OF ABSTRACTS	NUMBER OF PATENTS
BBJN	350	
BCJD	1	
BCJG	1	
BCJN	263	
BCPC	1	
BDJF	3	
BDJL	12	
BDJN	4	
BDPF	902	902
BDPL	30	30
BDPN	88	88
BFJN	323	
BGJF	51	
BGJM	1	
BGJN	721	
BGJR	1	
BGJV	4	
BGPV	1	1
BHJN	166	
BKJG	153	
BKJN	195	
BLJB	75	
BLJC	6	
BLJD	66	
BLJF	5	
BLJH	1	
BLJM	1	
BLJN	5	
BLJR	856	
BLJS	1	
BLJT	1	
BLJV	1	
BLJW	21	
BLTJ	1	
BMJG	2	
BMJR	240	
BMJW	18	
BMJX	1	
BMRJ	1	
BNJR	1	
BPJF	15	
BPJN	91	
BPPF	3	3
BSJG	8	2
BSJN	36	
BSPG	37	37
BTJF	103	
BTJG	49	
BTJH	1	
BTJL	5	
BTJN	16	

Figure 3

Example of Computer Printout Inventory of CA Abstracts by In-House Abstractor

IN HOUSE ABSTRACTORS BY CODE BY SECTION VOL 63 02/21/66 PAGE 01

CODE	NUMBER OF ABSTRACTS	NUMBER OF PATENTS
SECTION 1		
GLJR	1	
RCFQ	2	
RCTT	5	
RCVD	1	
RCVQ	1	
SFTT	1	
SGTT	5	
SHTT	1	
SNDC	1	
SNTT	52	
SXTT	1	
SZTT	2	
TCNG	1	
TCSL	1	
VNJZ	7	
SECTION TOTAL	82	
SECTION 2		
BDPN	1	1
BGJV	1	
BLJB	2	
BLJC	2	
BLJD	2	
BLJR	10	
BLJS	1	
BLTJ	1	
BMJR	23	
BSPG	1	1
BTJQ	2	
BWJY	1	
BYJY	53	
GBJR	6	
GHPF	1	1
GKJY	2	
GLJR	13	
GMJR	3	
GPJR	3	
GPJW	3	
GRJF	1	
GRJG	2	
GZJR	7	
HSJR	1	
MCCR	1	
MCRK	1	

Figure 4
Example of Inventory of CA Abstracts by In-House Abstractors, Arranged by
Section of CA

CHEMICAL ABSTRACTS SERVICE

CODEN BASED INVENTORY REPORT NUMBER 1 CA VOLUME 63
SOURCE OF ABSTRACTS

CODEN	NUMBER OF ABSTRACTS	PAGE 1	
AAAH	9		
AAAM	1		
AABC	8	7 BY CA	1 BY OTHER ABSTRACT JOURNALS
AABF	2		
AABI	33	29 BY CA	4 BY OTHER ABSTRACT JOURNALS
AABN	9		
AABQ	7		
*AABS	3	1 CREDITED TO RUSSIAN JOURNALS	
AACH	79		
AACI	7		
AACR	6		
AAFA	3	1 BY CA	2 BY OTHER ABSTRACT JOURNALS
AAFB	5		
AAFP	13		
AAFR	3		
AAGE	18		
AAGP	4		
AAGS	5		
AAGW	9		
AAHO	1		
AAKG	1		
AALS	1		
AALT	2	NONE BY CA	
AAHL	38		
AAHM	11		
AANE	6		
AANI	1		
AANL	41	37 BY CA	4 BY OTHER ABSTRACT JOURNALS
AAOS	1		
AAPL	2		
AAQA	33		
AARE	2		
AASC	3		
AASH	1		
AATF	9	1 BY CA	8 BY OTHER ABSTRACT JOURNALS
AATR	1		
AATS	3	1 BY CA	2 BY OTHER ABSTRACT JOURNALS
AAUP	1		
ABAB	19		
ABAH	9		
ABAI	1		
ABBI	184		
ABBS	3		
ABCH	102	100 BY CA	2 BY OTHER ABSTRACT JOURNALS
ABCL	18		
ABCO	6		
ABCZ	1		
ABEB	3		
ABEG	5		
ABEM	4	1 BY CA	3 BY OTHER ABSTRACT JOURNALS

* INDICATES SECONDARY JOURNALS

Figure 5
Example of Inventory of CA Abstracts by Source in Coden Order

CHEMICAL ABSTRACTS SERVICE										PAGE 1
CODEN BASED INVENTORY REPORT NUMBER 7MO					CA VOLUME 63	PAPER ABSTRACTS 85859				
CUMULATIVE NUMBER OF JOURNALS	CODEN	ABSTRACTS PER CODEN	CUMULATIVE NUMBER OF ABSTRACTS	PERCENT OF TOTAL	CUMULATIVE NUMBER OF PRIMARY ABSTRACTS	PERCENT OF TOTAL	CUMULATIVE NUMBER OF ONE SHOT JOURNALS	CUMULATIVE NUMBER OF ONE SHOT ABSTRACTS	PERCENT OF TOTAL	
1	DIAS	2073	2073		1		2073			
2	CORE	926 *	2999		2		2999			
3	NATU	918 *	3917		3		3917			
4	CHRE	789 *	4706		4		4706			
5	XAE*	789	5505	5.5	5	(5.5)	5505		0.0	
6	JOC	767	6272		6		6272			
7	BBAC	747 *	7019		7		7019			
8	QANK	717 *	7736		8		7736			
9	CHRE	716 *	8296		9		8296			
10	BLJD	439	8735		10		8735			
11	JPCB	437	9172	10.2	11	(10.2)	9172		0.0	
12	PHLT	432 *	9604		12		9604			
13	JBCB	422	10026		13		10026			
14	CHRE	422	10420		14		10420			
15	ZGAM	348	10768		15		10768			
16	FTV7	387 *	11195		16		11195			
17	XCTA	382 *	11547		17		11547			
18	BSCF	380	11957		18		11957			
19	TELE	365	12322		19		12322			
20	XRRR	340	12662		20		12662			
21	NAAC	340 *	13006	15.1	21	(15.1)	13006		0.0	
22	PSE8	339 *	13345		22		13345			
23	NGKZ	313 *	13658		23		13658			
24	CHRE	307	13967		24		13967			
25	CCOM	307	14274		25		14274			
26	ANCH	306	14580		26		14580			
27	IASK	305 *	14885		27		14885			
28	ICSJ	304	15189		28		15189			
29	CHRE	303 *	15492		29		15492			
30	ZPKH	301 *	15793		30		15793			
31	PHSS	301 *	16094		31		16094			
32	PHL7	299	16393		32		16393			
33	CCCC	299	16692		33		16692			
34	CHRE	299	16991		34		16991			
35	JAP1	287 *	17278	20.1	35	(20.1)	17278		0.0	
36	PRV8	285	17561		36		17561			
37	AECA	281 *	17842		37		17842			
38	CHRE	280	18122		38		18122			
39	CHRE	279	18401		39		18401			
40	CJCH	267	18662		40		18662			
41	VMSD	260 *	18922		41		18922			
42	JUPS	254 *	19176		42		19176			
43	CHRE	247 *	19423		43		19423			
44	CHRE	246 *	19669		44		19669			
45	ZORR	246	19916		45		19916			
46	TAIM	244 *	20160		46		20160			
47	MUCI	244	20404		47		20404			
48	IMOC	242	20646		48		20646			

* DENOTES SOME OR ALL ARE COPY ABSTRACTS ** DENOTES U.S. GOV'T PRIMARY JOURNAL OR R AND O REPORTS

Figure 6
Example of Inventory of CA Abstracts

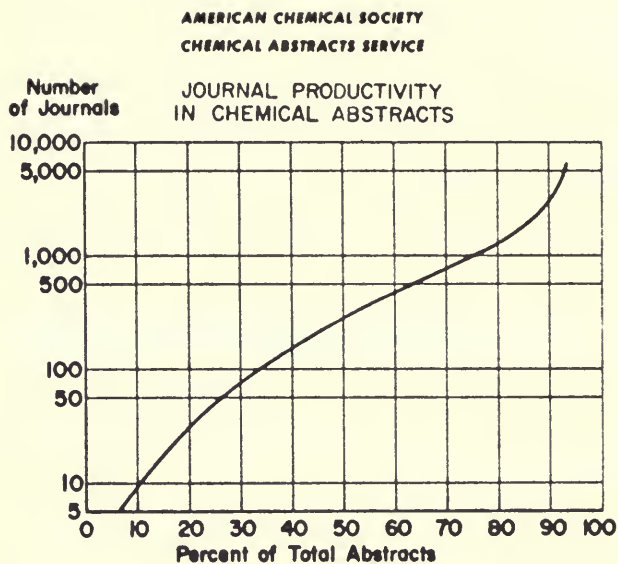


Figure 7
Four-Year Cumulation or Semi-Log Paper of Journal
Productivity of Abstracts

CHEMICAL ABSTRACTS SERVICE
CODEN BASED INVENTORY REPORT NUMBER 3 CA VOLUME 63
PRIMARY JOURNALS COVERED BY COPY ABSTRACTS

PAGE 1

CODEN	YR	VOL	ISS	PAGE	TITLE	CODEN	YR	VOL	ISS	PAGE
TOTAL UNIDENTIFIED PRIMARY JOURNALS 1001										
ANALIS OA ACADEMIA BRASILEIRA DE CIENCIAS										
AA6C	1962	34	1	181	199 \$01	CHZE	1964		4	389
	TOTAL									199 \$02
ANNALS OF APPLIED BIOLOGY										
AA81	1963	52		327	\$01	CHZE	1965			1468
AA81	1964	53		291	\$01	CHZE	1965		4	1468
AA81	1965	53		291	\$01	CHZE	1965			2200
AA81	1966	53		291	\$01	CHZE	1965			2333
AA81	1964	53	4		309 \$01	CHZE	1965		1	1269
	TOTAL									\$02
ANALE. INSTITUTUL CENTRAL DE CERCEATARI AGRICOLE. CONSILIUL										
ANALIZELOR SI REZULTATELOR ALE CERCETARILOR IN DOMENIUL										
AMELIOARE, PROTECTIA PLANTELOR SI TEHNOLIGIE AGRICOLA.										
AAFA	1961	29	235		\$01	CHZE	1964		4	2069
AAFA	1961	29	9		\$01	CHZE	1964		4	2048
	TOTAL		2							\$02
ARCHIVIO...DE VEECHI... PER L/ANATOMIA PATOLOGICA E LA										
MEDICINA CLINICA										
AALT	1961	35	3	729	740 \$01	81A8	1965	46		894
AALT	1961	36	2	345	382 \$01	81A8	1965	46		582
	TOTAL		2							\$02
ATTI DELLA ACCADEMIA NAZIONALE DEI LINCEI. RENDICONTI,										
CLASSE DI SCIENZE FISICHE, MATEMATICHE E NATURALI										
AANL	1961	30	437		\$01	CHZE	1964		4	598
AANL	1962	32	567		\$01	CHZE	1965			1697
AANL	1962	96	805		\$01	CHZE	1964		4	607
AANL	1962		96	517	533 \$01	CHZE	1964		4	127
	TOTAL		4							\$02
ATTI DELLA ACCADEMIA DELLE SCIENZE DI TORINO. CLASSE DI										
SCIENZE FISICHE, MATEMATICHE E NATURALI										
AATF	1963	97	258		264 \$01	CHZE	1965		1	637
AATF	1962	96	296		096 \$01	CHZE	1964		2	127
AATF	1962	96	903		\$01	CHZE	1964		4	259
AATF	1962	96	534		\$01	CHZE	1964		4	688
AATF	1963	97	757		765 \$01	CHZE	1965		1	648
AATF	1963	97	635		648 \$01	CHZE	1965		1	713
	TOTAL									\$02

Figure 8
Example of Print-out Report of Inventory of Copy Abstracts Used in CA, Arranged by Primary Journal

(TITLE)				
_____	_____	_____	_____	_____
(YR)	(VOL)	(ISSUE)	(PAGE)	
_____	_____	_____	_____	_____
(CA VOL)	(CA ISSUE)	(SECT)	(COLUMN)	(FRAC) (SUB F)
_____		_____		
(ABSTRACT NO.)		(ABTRACTOR)		
_____		_____		_____
(REQUESTED BY)		(DATE)	(OUT TO)	(DATE)
NOT YET REC'D ☐		NOT REC'D ☐		
NOT ON SHELF - NOT CHECKED OUT TO STAFF ☐				
VOL	SECT.	NO.	PAGE	

CAS CODEN DIRECTORY. TITLE SEQUENCE		PAGE	I
CODEN	TITLE	SEQ. NO.	
AGHY	ABHANDLUNGEN AUS DEM GESAMTGEBIETE DER HYGIENE.	000385	01
AMEH	ABHANDLUNGEN AUS DEM INSTITUT FÜR METALLHÜTTENWESEN UND ELEKTROMETALLURGIE DER TECHNISCHEN HOCHSCHULE, AACHEN.	000425	01 000425 02
APWP	ABHANDLUNGEN DER PREUSSISCHEN AKADEMIE DER WISSENSCHAFTEN, PHYSIKALISCH-MATHEMATISCHE KLASSE	000570	01 000570 02
ABMB	A.B.M. BOLETIM (ASSOCIACAO BRASILEIRA DE METAIS (SAO PAULO)) A.B.M. BOL. ASSOC. BRASIL. METAIS (SAO PAULO)	000700T01	000700T01
AVCM	ABRASIVE AND CLEANING METHODS.	000730	01
APKR	ABRIDGED SCIENTIFIC PUBLICATIONS FROM THE KODAK RESEARCH LABORATORIES.	000765	01 000765 02
ATPM	ABSTRACTS OF SCIENTIFIC AND TECHNICAL PUBLICATIONS FROM THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY.	001095	01 001095 02
ARBB	ACADEMIA REPUBLICII POPULARE ROMINE, INSTITUTUL DE BIOCHIMIE, STUDII SI CERCETARI DE BIOCHIMIE	001960	01 001960 02
ACCF	ACCION FARMACEUTICA	002690	01
AGNR	ACQUISITIONS MEDICALES RECENTES. JOURNEES MEDICALES ANNUELLES DE BROUSSAIS - LA CHARITE.	002930	01 002930 02
ABIM	ACTA BIOLOGICA ET MEDICA GERMANICA, SUPPLEMENTBAND	003630	01
ACUJ	ACTA ET COMMENTATIONES IMPERIALIS UNIVERSITATIS JURIEVENSIS (OORPATENSIS)	004045	01 004045 02
ATUA	ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS (OORPATENSIS) A.* MATHEMATICA, PHYSICA, MEDICA	004050	01 004050 02
ATUC	ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS (OORPATENSIS) C.* ANNALES	004055	01 004055 02
ALSM	ACTA LEIOENSIA. SCHOLAE MEDICINAE TROPICAE	004850	01
ALHM	ACTA LITTERARUM AC SCIENTIARUM REGIAE UNIVERSITATIS HUNGARICAE FRANCISCO-JOSEPHINAE, SECTIO MEDICORUM	004890	01 004890 02
AMTO	ACTA NIPPONICA MEDICINAE TROPICALIS.	005770	01
APAI	ACTA PHYSIOLOGICA ACADEMIAE SCIENTIARUM HUNGARICAE, SUPPLEMENTUM	006530	01 006530 02
APYP	ACTA PHYSIOLOGICA POLONICA LITTERAE SOCIETATIS PHYSIOLOGORUM POLONORUM ACTA PHYSIOL. POLON.	006580T01	006580 02 006580T01
APXG	ACTA PHYTOTAXONOMICA ET GEOBOTANICA.	006670	01
APOP	ACTA PSYCHIATRICA Y PSICOLOGICA DE AMERICA LATINA	007010	01

Figure 11
Sample Page of Title Sequence of CAS Coden Directory

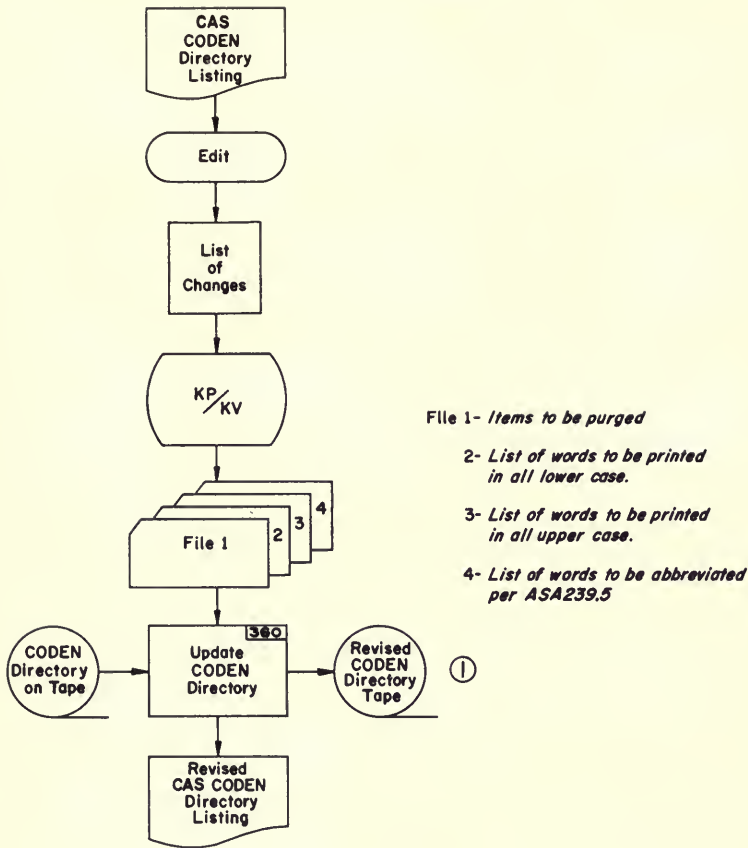


Figure 12
Flow Chart for Revision of Present CAS Coden Directory

File Generation-CODEN Additions

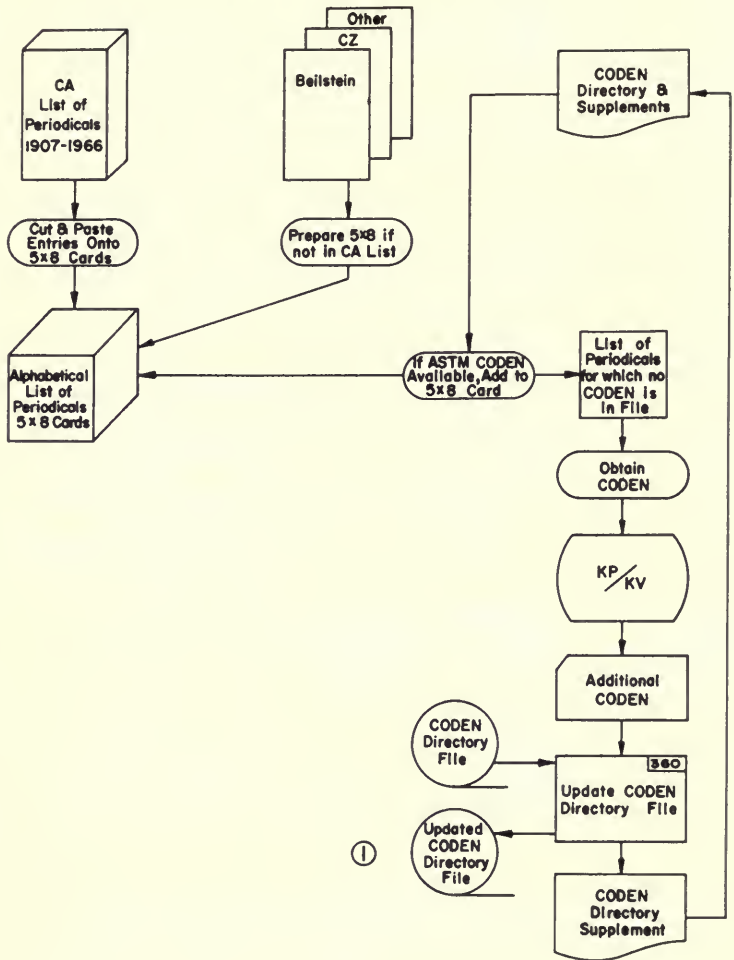


Figure 13
Flow Chart for Preparation of Master File of Periodicals with Coden

Checking Edition Preparation

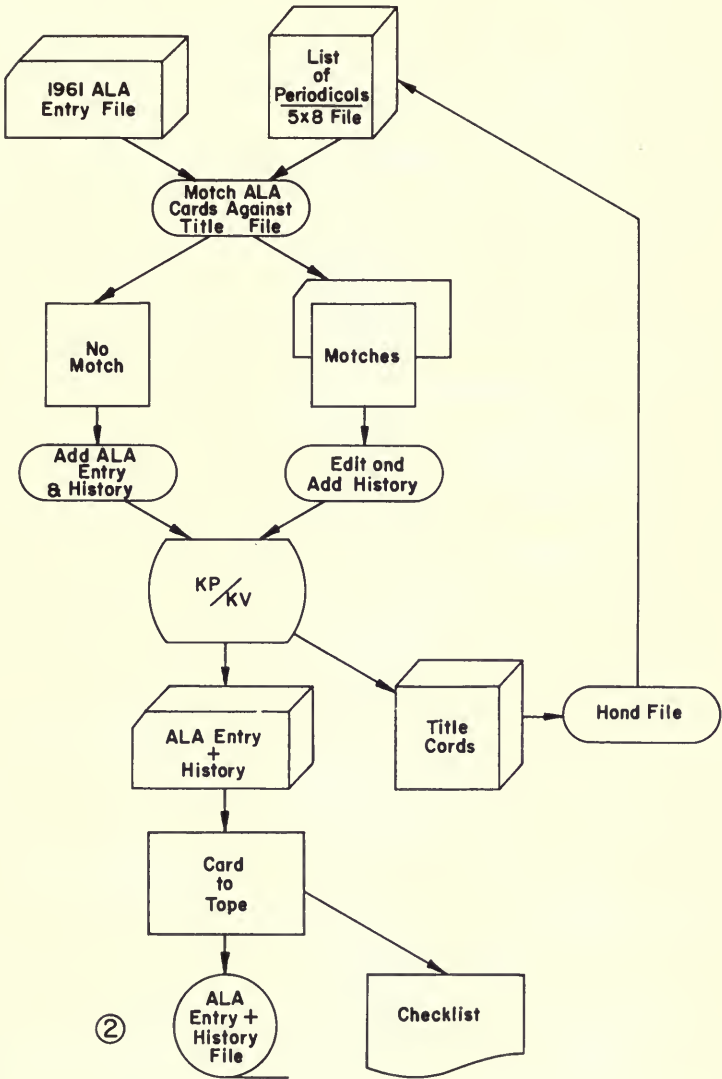


Figure 14
Flow Chart for Preparation of Checking Edition of List of Periodicals

Library Holdings Tabulation

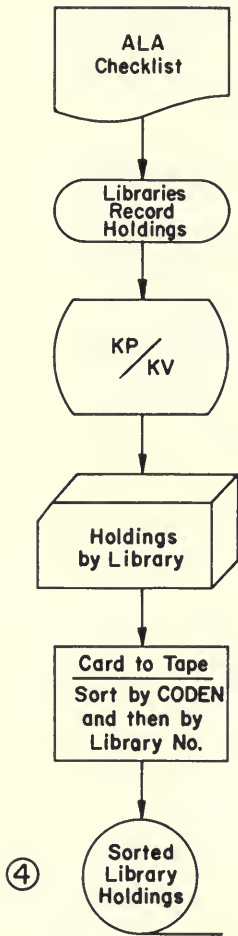


Figure 15
Flow Chart for Adding Library
Holdings of Periodicals to
Computer Record

Work Flow for Periodicals List

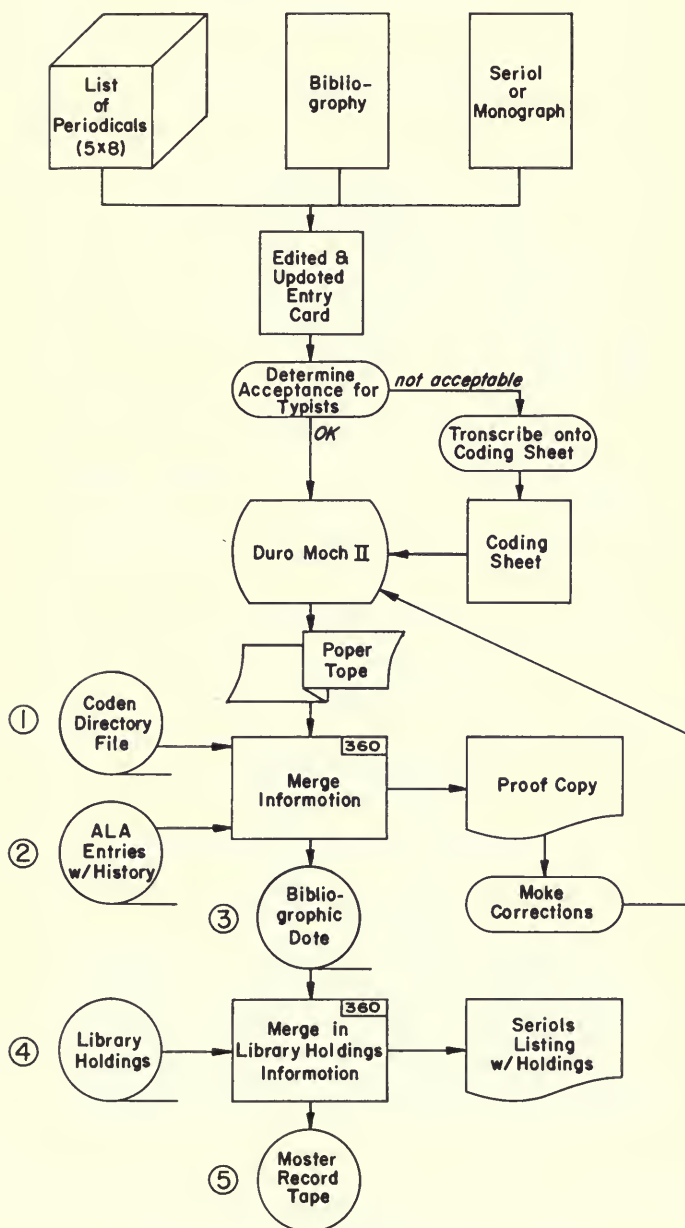


Figure 16
Flow Chart for the Preparation on Tape of the Final Complete List of Periodicals

Prepare Specialized Listings

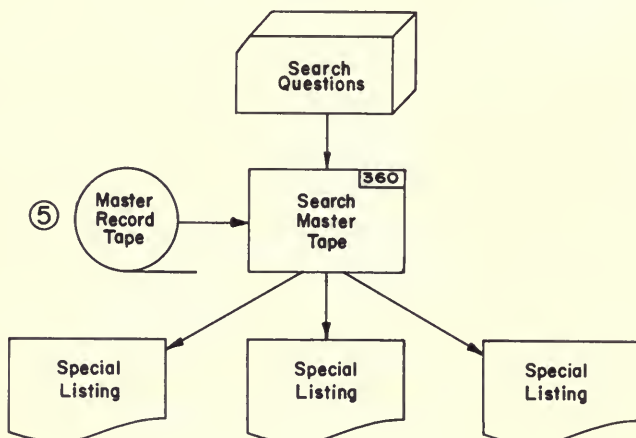


Figure 17
Flow Chart for the Preparation of Special Lists of Periodicals

CHEMICAL TITLES

KEYWORD-IN-CONTEXT INDEX

WITH AROMATIC ALDEHYDES, AND THE GEOMETRICAL
 WITH ALIPHATIC ALDEHYDES, MALONIC ESTERS, AND
 OF SETTING FREE THE ALDEHYDIC GROUP IN PYRIDOXAL.
 OF ALDOSES AND THEIR ALDEHYDRO DERIVATIVES WITH COMPOUN
 ESTER OF THE DIOLS- ALDER ADDUCT OF PETRO SELINIC ACID
 CYCLO PENTADIENOL-ALDE- ALDER REACTIONS WITH TETRA CHLORO
 AND ORALLY ADMINISTERED ALDO STERONE. EVIDENCE FOR EXTRA
 RIBOSE-5-PHOSPHATE ALDOOLASE IN THE SYNTHESIS OF DEDXY
 OF METHYL DOPA (ALDOMET) ON CARDIOVASCULAR RESPON
 TIVES + CONDENSATION OF ALDOSES AND THEIR ALDEHYDRO DERIVA
 SYNTHESIS OF 2,2- ALDOSYLIDENE-BIS-(5,5-DI-METHYL-1,
 HEXANEDIONE) AND 2,2- ALDOSYLIDENE-BIS-(5,5-DI-METHYL-1,
 PHOTOMETRIC ANALYSIS OF ALDRIN AND DIELDRIN IN A TECHNICAL
 RING THE MANUFACTURE OF ALDRIN. + COMPOUNDS OBTAINED DU
 NTHOSINE FROM TRIFOLIUM ALFALFANDRUM SEEDS.

JS00-A-1966-0283
 BSCF-A-1965-3596
 JUAR-A-DD08-D129
 JPMS-A-DD55-DD28
 JAOC-A-DD43-DD33
 TELE-A-1966-0201
 JCIN-A-DD45-D264
 88RC-A-DD22-DD62
 AIPT-A-D159-D215
 JPMS-A-DD55-DD28
 JPMS-A-DD55-DD28
 JPMS-A-DD55-DD28
 CANW-A-DD10-1253
 MGKF-A-DD72-DD44
 XA CUSC-A-DD35-DD11

Figure 18
Sample of KWIC Index to Chemical Titles

CHEMICAL TITLES

BIBLIOGRAPHY

J. Clin. Invest., 45, No. 2 (1966)

JCIN-A-0045-0228	ORTE O WILLIAMS RM EFFECT OF EPINEPHRINE ON IMMUNE REACTIVE INSULIN LEVELS IN MAN. *	GRABER AL KUZUYA T 228-36
JCIN-A-0045-0249	GROSS RT PYRIDINE NUCLEOTIDES IN ERYTHROCYTE METABOLISM. *	SCHROEDER EA GABRIO BW 249-55
→ JCIN-A-0045-0264	BLEESOE T ISLAND DP COMPARATIVE FATES OF INTRAVENOUSLY AND ORALLY ADMINISTERED ALDO STERONE. * EVIDENCE FOR EXTRA HEPATIC FORMATION OF ACID HYDROLYZABLE CONJUGATE IN MAN. 264-9	LIOOLE GM BLOOMFIELD O SINCLAIR-SMITH P 269-55
JCIN-A-0045-0270	REGAN TJ OLDENHUTEL HA ACUTE METABOLIC AND HEMODYNAMIC RESPONSES OF THE LEFT VENTRICLE TO ETHANOL. *	KORDENHOIS G LEHMAN PH MOSCHOS CB 270-80

Figure 19
Sample of Bibliography Section of Chemical Titles

CHEMICAL TITLES

AUTHOR INDEX

→ BLECHTA Z	CHPU-A-0016-0041
BLEESOE T	JCIN-A-0045-0264
BLEIDNER WE	JPET-A-0150-0484
BLIEGEN HR	PHLT-A-0019-0712
	PHLT-A-0019-0708
	ELCA-A-0011-0093
BLOCH O	
BLOEMBERGEN N	PRLT-A-0016-0081
BLOMQUIST AT	JOCE-A-0031-0338
BLOODWORTH AJ	
	JSOD-A-1966-0299
BLOOM H	AJCH-A-0018-2039
BLOOMFIELD C	JSFA-A-0017-0039
→ BLOOMFIELD O	JCIN-A-0045-0264
BLOSSEY RG	TAMM-A-0236-0028

Figure 20
Sample of Author Index to Chemical Titles

CHEMICAL-BIOLOGICAL ACTIVITIES

ILLUSTRATIVE DIGEST

J. Biol. Chem., 241, No. 2 (1966)

259 STUDIES ON THE TRIPHOSPHOPYRIDINE NUCLEOTIDE-CYTOCHROME *f* REDUCTASE OF CHLOROPLASTS.

JBCW-A, 241, 2, 66, 279-85

Zanetti G, Forti G; Univ. Milan.

The kinetic properties of a spinach LEAF CHLOROPLAST crystalline Flavoprotein enzyme, both as

TPNH-cytochrome *f* reductase and as

TPNH diaphorase, as well as its molecular weight, content of

146145 FAD, and reductive inactivation, were studied.

259-1 Kinatic reaction studies of the activity of TPNH-cytochrome *f* reductase gave a K_m for the coenzyme of 2.1

53576 micromolar, suggesting the formation of a ternary complex of

Cytochrome *f*, and the enzyme.

259-2 The REDUCTION of Ferricyanide by the enzyme gave a K_m of 30.7 micromolar and involved the reaction of

53558 TPNH with the enzyme to form

TPN and reduced enzyme followed by the reaction of the reduced enzyme

to form

Ferrocyanide and oxidized enzyme.

259-3 The DIAPHORASE ACTIVITY of the enzyme was reversibly inactivated by preincubation with

TPNH;

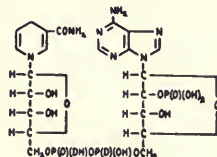
TPN prevented and reversed this INACTIVATION.

259-4 57136 Urea irreversibly inactivated DIAPHORASE ACTIVITY of the enzyme and converted the reversible inactivation by

TPNH into an irreversible process.

259-5 The activity of TPNH-cytochrome *f* reductase was not affected.

259-6 The functioning of FAD and the maintenance of bonds sensitive to Urea are essential to the preservation of the native structure of the enzyme required for activity.



53576

Figure 21
Sample Page of the Digest Section of CBAC

CHEMICAL-BIOLOGICAL ACTIVITIES

KEYWORD-IN-CONTEXT INDEX

→ 253-7	MITOCHONDRIA fractions+ When the	CHLOROPLAST end MITOCHONDRIA fra
259	properties of a spinach LEAF	CHLOROPLAST crystalline Fieovpro
387-3	PHOSPHORYLATION decreased to zero in	CHLOROPLAST fragments from LEAVE
38	."	CHLOROPLAST STRUCTURE AND FUNCTI
387-6	STRUCTURAL INTEGRITY of the	CHLOROPLAST, permitting the fine
8	ADP PHOTOREDUCTION by isolated	CHLOROPLASTS by 200-350%." + p
288-1	factor was removed from the	CHLOROPLASTS by washing or hypot
383	bonate, isolated whole spinach	CHLOROPLASTS formed a small emou
38-10	distinguishable from those+	CHLOROPLASTS from CALLUS pest th
431	RELEASED PHOTOPHOSPHORYLATION BY	CHLOROPLASTS FROM CHINESE CABBAG
431	se cabbage, in comparison with	CHLOROPLASTS from control LEAVES
431-4	Y MECHANISM that would prevent	CHLOROPLASTS from functioning ei
38-9	d in comparison with+ The	CHLOROPLASTS from growing CALLUS
387-1	Pyridine nucleotides in the	CHLOROPLASTS from immediately de
387-1	at 80 and 20%, respectively,+	CHLOROPLASTS from LEAVES floated

Figure 22
Sample of KWIC Index to CBAC

CHEMICAL-BIOLOGICAL ACTIVITIES

MOLECULAR FORMULA INDEX

DIGEST NUMBER	MOLECULAR FORMULA	REGISTRY NUMBER
5	$C_{21}H_{30}N_7$ $O_{17}P_3$	53576
20-4		
22-5		
77-11		
→ 259-1		
297-3		
411		
82	$C_{21}H_{30}O_2$	57830
166-1		

Figure 23
Sample of Molecular Formula Index to CBAC

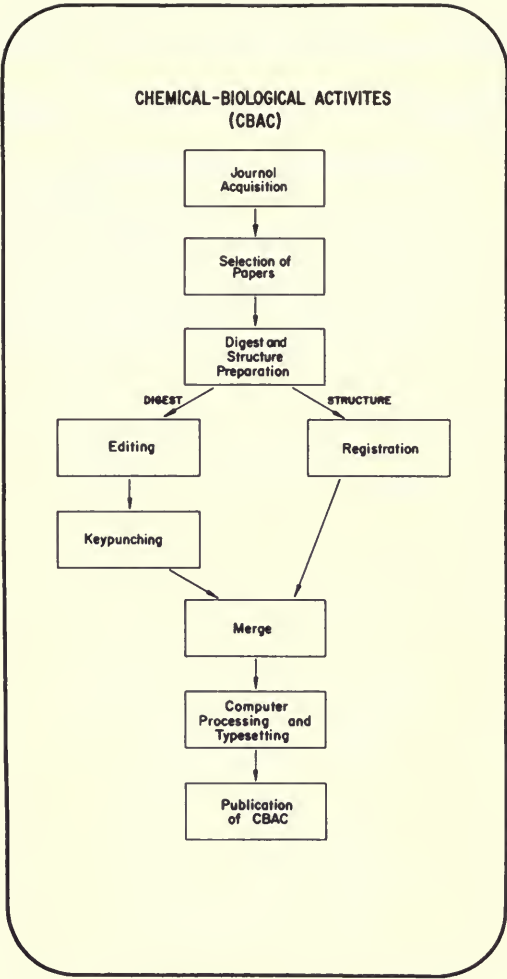


Figure 24
Flow Chart for the Preparation of CBAC

Master Random Access File

Auxiliary Information Files

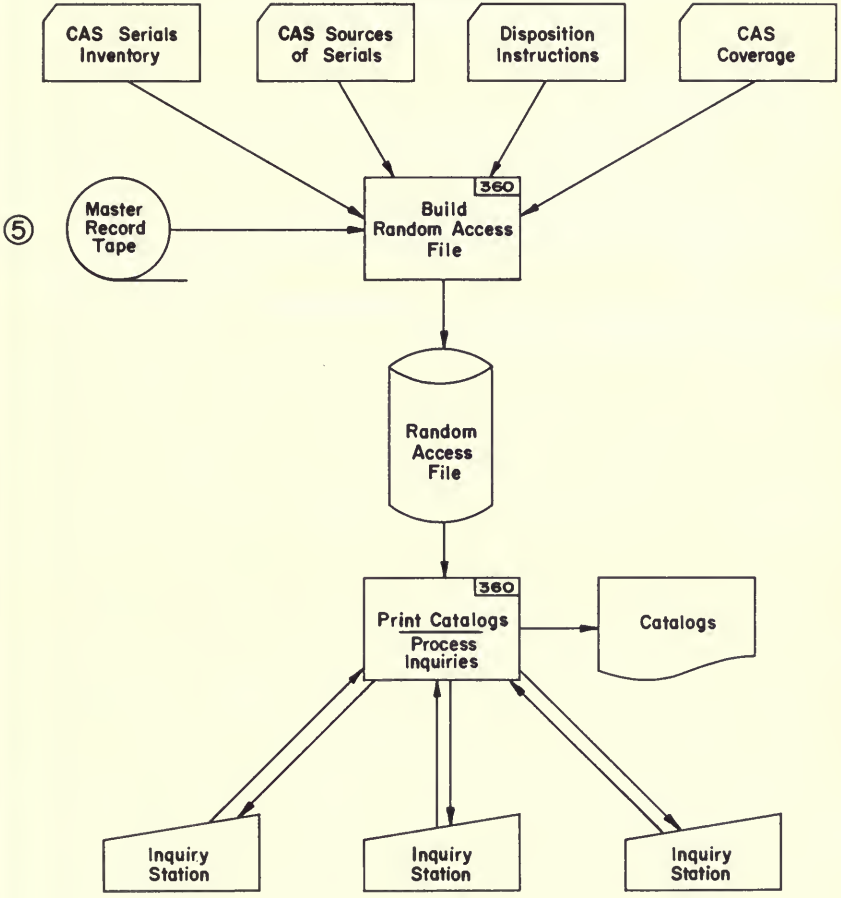


Figure 25
Flow Chart for Input and Output of Master Random Access File

CAS Computer Configurations

1. IBM 360 Model 40
 - 32 K storage
 - 4 7-channel tape drives (30 KB tapes)
 - 1 9-channel tape drive
 - 2 printers
2. IBM 7010
 - 80 K storage
 - 8 7-channel tape drives
 - 1301 disk file
3. Two remote terminals—IBM 1050's
4. Increase in storage capability of 360 Model 40 to 128 K—
March 15, 1966
5. Full operating system for the 360 Model 40, disk file and
floating point arithmetic—April 1966

DATA PROCESSING IN THE TEXAS A & M UNIVERSITY LIBRARY

Bruce W. Stewart

The Texas A & M University Library embraced automation as a way of life when it became the first library in the Southwest to employ a Data Processing Supervisor as a full-time Library staff member in September, 1964. The creation of such a position as part of the Library staff was only one of several favorable circumstances which combined to provide the necessary foundation for the achievements outlined in this paper.

In addition to an enthusiastic University administration which provided requested supplemental funds for a special conversion project, the Library has access to the University's centralized data processing facility, which is one of the largest such University installations in the Southwest. The Data Processing Center houses an IBM 7094-1401 computer system with 14 magnetic tape drives, two separate off-line 1401 tape systems (one with a 1404 printer), and a battery of high speed sorters, collators, and card punches. This tremendous hardware capability has proved to be a great asset to our automation program.

A second asset of a more subjective nature is the Library's membership in a co-operative effort involving the libraries of Texas A & M, Rice University, and the University of Houston. While ninety miles separate College Station from the University of Houston and Rice in Houston, the association has proved very beneficial to each library. In the fall of 1964 the three libraries met to discuss their mutual interest in automation. Since available computer equipment varied tremendously with each University, an agreement was reached to share ideas and co-operative effort at the systems and design level, with meetings to be arranged roughly three times per year. Each library reviewed its most pressing needs, and the final result was that the University of Houston Library decided to tackle the acquisitions operation, Rice University Library was already preparing to automate its circulation operation, and Texas A & M elected to attack the serials control problem.

Bruce W. Stewart is Coordinator of Data Processing for The Library, Texas A & M University.

Probably the most outstanding result of the co-operative agreement, in addition to the strong personal relationships that developed, was that Texas A & M was able to justify and implement an automated circulation system in an extremely short time due to the sharing of both experience and advice of the Rice Library. Although the two circulation systems are related only by the IBM 357 Data Collection Units used by each, considerable time and effort were saved by Rice's generosity. With implementation of Texas A & M's Automated Serials Control operation, the University hopes to be in a position to repay its debt to the co-operating institutions.

The Texas A & M University Library is wholeheartedly involved in data processing applications and the automation of clerical operations. While not concerned immediately with the more sophisticated area of information retrieval, the Library looks forward to the time when such systems will become a reality. Although the two applications just mentioned, serials and circulation, represent the two major areas of automation, several other routine functions also utilize punched cards. The Library's physical property inventory is maintained with a punched card system which allows rapid reconciliation with the University's inventory records at the close of each fiscal year. The Acquisitions Department also maintains a punched card file by purchase order number, coded to allow rapid breakdown of outstanding encumbrances by categories such as books, serials, and others. Finally, Acquisitions runs an annual analysis of vendor performance (primarily the state contract jobber which changes from year to year) by punching and analyzing on an IBM 1401, data which provides the percentage of books delivered by five day increments and the percentage of books delivered for each discount received.

Automated Circulation System

In February 1965, the University approved a supplemental grant to the Library in response to a proposal which outlined and justified the conversion necessary to implement an automated circulation system. There were two major steps involved in the initial conversion of our collection for the automated system, viz., keypunching, and gluing pockets and inserting cards.

Of the Library's half-million volumes, it was necessary to produce IBM bookcards for a circulating collection of approximately 200,000 volumes. This represents the circulating collection in the main library only; branch libraries were not included in the conversion or in the plans for automation. The branches will be absorbed into a new \$3,800,000 library building to be completed in 1968; at that time the necessary IBM bookcards will be keypunched in another

crash project. In addition, since bound serial holdings are not circulated, it was unnecessary to produce a machine readable bookcard for the additional 200,000 bound serial volumes in the main library building.

After several methods of producing the necessary machine readable identification for each circulating volume were investigated, local keypunching under the Library's own management was selected as the most practicable approach. Alternate methods considered included optical scanning of pencil coded sheets and contract conversion by a private firm. Neither of these possibilities was acceptable, however, either because of simple unavailability of the service desired or because of the problem of shipping irreplaceable shelflist trays to and from the contract keypunching firm ninety miles away. The deciding factor in the final decision was the availability of the University Data Processing Center's keypunches on a second shift basis at no cost to the Library.

Five girls were hired at \$1.25 per hour to work 4-1/2 hours each weekday night, 5:30 p.m. to 10:00 p.m., plus 3-1/2 hours on Saturday morning. The girls read directly from the shelflist cards in their catalog trays, which were carried to the Data Processing Center and back each evening. The girls typed the accession number and call number for each bound volume except for those books which because of classification number or shelf card notation were known to be non-circulating, entering an average of 25 keystrokes per card. Figure 1 illustrates the full bookcard format. Only one of the girls was an experienced keypunch operator although all were capable typists. By the end of the first week they were averaging 1,000 cards per night per girl and maintained this average for approximately seven weeks, the time required to complete the project. Card output was limited not by punching speed but by the time required for the girls to interpret each shelflist card according to prescribed instructions.

The keypunched cards were kept in shelflist order and periodically listed on an IBM 1401. The listings were then sight-verified against the original shelflist cards by two additional girls employed full-time during the day for the duration of the project. Error cards, omitted cards, and cards for non-circulating volumes were pulled or repunched. All original cards, which were punched in regular blank card stock, were then input to an IBM 1401 program which scanned the fields, adding fill periods and 357 control punches, and then punched the final card into the Library's Master Bookcard stock. This program also printed each volume's accession number on a gummed 1/4" by 1-1/4" label.

The final master cards, kept in shelflist order in labeled boxes with the corresponding gummed labels, were then run through an IBM 557 interpreter. Costs for keypunching and verifying were as follows:

Wages

Keypunching: March 22 - May 11, 1965

Sight verifying: April 5 - May 28, 1965 \$1800

Supplies

200,000 Master Bookcards \$ 410

217,000 Stock Tab Cards \$ 255

200,000 Pocket Labels \$ 260

Non-billed Operations

Listing on IBM 1401 for verification 15 hours

Reproducing for Master Card on IBM

1401 20 hours

Interpreting on IBM 557 40 hours

The book collection lacked pockets which were glued in by high school students at \$.75 per hour. The boys worked from 8 a.m. to 5 p.m. in unairconditioned stacks during the summer months. Working in teams of two, ten boys averaged 7500 books per day. The gluing operation was virtually complete in five weeks.

These students plus an additional five people were also utilized to match the Master Bookcard to the book, insert it in the pocket, and attach the accession number label to the pocket. The insertion operation was in simultaneous operation by the second week in June and continued to August 20. Most of September was also spent with one regular library staff member working full-time on the backlog of snags.

Costs for the gluing and insertion operation were as follows:

Wages

Gluing Pockets: June 1 - July 7, 1965

Match and Insert: June 15 - August 20, 1965 \$4600

Supplies

200,000 Bookpockets \$ 991

5 Gluing Machines \$ 55

30 gallons glue \$ 105

Roughly, the total conversion cost was approximately 3.9¢ per book and the project, start to finish, took five months.

In addition to the conversion of the collection it was necessary for each Library patron to have a plastic card with borrower identification in machine readable form punched into it. The Registrar's office already provided a student identification card, which was modified to IBM 357 system specifications. This card is custom laminated by the University and embossed in the Registrar's office with student

permanent number, name, and birth date. Each currently enrolled student receives a new card at the beginning of each fall semester, which provides automatic yearly validation.

All other non-student users, including University faculty and staff, faculty and staff family, and local area and special users are issued a permanent factory-produced card purchased by the Library. These are non-embossed cards on which the name and user number are handwritten. All non-student users are required to provide their Social Security number as identification. A library user thus either presents his student ID card to be punched or requests a Library Card the first time he enters the Library during the year. The patron's name, mailing address, and identification number are recorded for keypunching and addition to a magnetic tape file of names and addresses. The punched ID card is immediately returned, and the patron is ready to charge his book. The six digit student identification number is recorded with three trailing zeros, and an additional prefix number is assigned to both Social Security and student numbers to indicate the status of the borrower for later analysis of Library usage.

The data processing equipment in the Library consists of an 082 Sorter, an 085 Collator, an 026 Key punch, an 013 Badge Punch, and the IBM 357 Data Collection System. The IBM 357 system includes two model six input stations with cartridge manual entry, mounted at the Circulation Desk on either side of a 358 control unit and receiving keypunch which are enclosed in a soundproof plywood box built into the desk. The regular 026 Key punch services the needs of other Library departments upon request as well as all regular processing of new acquisitions and other Circulation Department key-punching.

A brief description of operation of the IBM 357 system is as follows: A patron presents a book to be checked out with a valid current identification card. The Circulation clerk selects the correct date due cartridge to be used for the charge and inserts the bookcard, identification card, and cartridge to activate the IBM 357 input station. The output keypunch produces two cards, each containing book identification, borrower identification, and date due (see Figure 2). The two cards are uniquely coded in cc 80. The first card is placed in the pocket with the bookcard as proof of charge and as notification to the borrower of date due. The second card remains on the key-punch stacker, and the accumulated charge cards are input to the nightly computer processing run.

As books are returned, they are held on a booktruck with their corresponding return cards. Immediately before the truck is released for shelving, the return cards are sorted into accession number order and matched with the collator against a deck containing one card (see

Figure 3), for each book on Hold. The Hold cards are likewise maintained in accession number order and contain only the accession number in machine readable form. Any matching cards that merge flag books that are to be removed from the truck before it is released to the stacks. While a low-volume operation, the sorter and collator in the Library are necessary even in a computer-based system such as this one, not only for Circulation processing but for various other IBM card filing routines that can be done more conveniently on a mechanical basis.

A manual back-up charging procedure is available in the event the IBM 357 system breaks down, or in the event a book is presented at the Circulation Desk with the Master Bookcard missing or mutilated. The procedure utilizes a three-part IBM card-size form (see Figure 4), which is used manually to record book identification, borrower number, and date due. One part then goes with the book in place of the Return Card, one part is filed as a temporary record of charge, and the third part is sent to the keypunch operator to have the required Master Card or Charge and Return Card set keypunched. The Charge Card thus produced is entered into the day's accumulation of charge cards and processed normally by the computer. The Master Bookcard or Return Card is filed manually to await the return of the book. All returns containing the temporary return form must have the Bookcard or Return Card waiting in the file pulled, and then must be processed through the normal Hold procedure.

The Library believes that a book should either be in its proper place on the shelf, or it should be reflected by a charge to a person or a location. Location charges are achieved by some fifteen special plastic charge cards kept in the Circulation Office which are punched with special numbers for "Reserve," "Bindery," "Interlibrary Loan," etc. All charges are maintained in one circulation file in call number sequence so that location and patron charges are interfiled for easy reference on the circulation listing.

One additional feature of the system is worth special note. The use of an accession or item number unique to each volume permits a thief-proof door check system. In a "two-card" system such as this, all information is machine recorded: book identification, borrower identification, and date due. No manual stamping or validating is required. The accession number is stamped both in the book and on a pocket during initial processing, and this pocket is then glued firmly in the book. The accession number labels which were computer printed during the conversion period also provide the item number for back holdings. The Library Doorman thus matches item number on the pink Return Card and item number on the pocket as proof of valid charge. Other information, including date due, need not be checked. The unique item number guarantees the current validity of the charge and also prevents the common dodge of charging a book

and then using the Return Card in a second book to get it past the door check.

Regular nightly operation, which has continued without interruption since October 1965, proceeds as follows. The Circulation Department delivers to the Data Processing Center by 10:30 each night except Sunday the accumulated Charge and Return cards in random order. These updates are sorted, edited, and processed by the IBM 7094, and a new listing of books in circulation (see Figure 5) is printed on a two part paper. Error messages, daily statistics, and a listing of extreme overdues immediately follow on the listing (see Figure 6). First and second overdue notices and fine notices (see Figure 7) are printed on continuous form postcards with addresses ready for mailing. The update cards and printed outputs are returned to the Library by 8:15 each morning. The two copies of the circulation listing are bound unburst and placed on the desk for patron use, and the postcards are separated and mailed. One copy of the error messages and statistics is retained by Circulation for further action, and the second copy is forwarded to the Data Processing Supervisor for inspection.

Currently, there are about 10,000 charge records on tape and probably an average of 700 updates (charges and returns) for each day. The nightly run requires less than three minutes of IBM 7094 time, plus ten to fifteen minutes of off-line printing on a 600 lpm IBM 1403. The program is written entirely in COBOL except for an assembly language core sort library subroutine. The program was written and checked out in approximately one month's time early in the fall of 1965. The most recent modification, however, was made in February 1966; it seems that odd situations which were not provided for in the program will continue to occur.

The total cost for rental of the data processing equipment located in the Library, less various standard educational discounts, is \$355 per month. A flat charge of \$400 per month is paid to Data Processing for use of machine time, resulting in a yearly cost of \$9,060. The continuous forms for the printed listing and the postcard stock, plus the other special card forms needed, result in an additional cost of approximately \$1,000 per year.

While the automated circulation system can be justified by virtue of performing at least the work of the approximately three and one-half clerks represented, the advantages of the mechanized operation over any traditional manual system cannot be over-emphasized. The system now provides better control than was ever possible under the old manual system. At the same time, it accommodates patrons by providing rapid, easy charging of books and by keeping an up-to-date, error free record of the location of books charged out.

Elimination of manual handling and filing of records as well as the tremendous saving of time effected by automatic processing of

overdue charges and delinquent returns has freed the circulation staff to assume additional duties while still maintaining almost immediate reshelving of books as they are returned. Perhaps most significant of all is the fact that the present system can accommodate double or triple the present volume of circulation with only a minimum increase in personnel. The automated system also offers great promise in aiding investigation of library usage at Texas A & M by analysis of accumulated charge data in conjunction with the Registrar's machine readable student records.

Automated Serials Control

The Library's decision to attempt automation of the Serials Record operation was made early in 1965. At that time a report outlining basic concepts of the system was prepared, based on the results of a Master's thesis investigation.¹ Initial programming and experimental operation of this proposed system were delayed, however, until January 1966 because of the higher priority assigned to the completion of the design and implementation of the automated circulation system just described. When completely operational, the automated serials system will perform three major tasks: (1) ordering and renewing subscriptions and recording payment for each, (2) recording arrivals and flagging delinquent arrivals, as well as keeping track of all serial holdings to date in the collection, and (3) providing the information required for more efficient control and analysis of operations. In addition, the data in the computer's records can be used to produce listings in various formats for use in public service areas. Such a mundane use alone of the stored data can almost justify the conversion to machine readable form.

Basic Concepts. The basic design concept of the data processing system which supports the serials operation is in some aspects a departure from traditional approaches. Historically, data processing techniques have been developed to accommodate what was considered a basic hardware limitation. This limitation may be described as the necessity for conserving data storage and processing time at nearly any cost. This concept dates from the early years of card oriented computer and unit record equipment systems. In the past, relatively slow interval operating speeds and relatively slow access to mass data storage dictated heavy coding and extreme abbreviation of data for applications involving extensive computer processing. The result has been data processing systems which in reality were designed to accommodate the machines involved and not the personnel who must ultimately handle and analyze the information produced.

Current sophistication of computer hardware and the promise of additional technological advance dictate that completely efficient hardware utilization is no longer the prime measure of effectiveness. Modern high speed computers with practically unlimited rapid access data storage capability allow feasible design of data processing systems suitable to the needs of the application as well as the personnel involved. While relatively inefficient internal processing procedures may be called for, the resulting performance of the system as a whole is the criterion that must be paramount. Input and output of information and data handling procedures are, therefore, organized to accommodate the library personnel involved. Data are stored in the computer uncoded and of variable length wherever possible, especially when coding and decoding would become cumbersome for clerical personnel. Stored in this form, the data can also logically be utilized by or converted to the format required for nearly any conceivable processing system.

The Serials Department personnel will be involved as little as possible in data processing lore or systems of special abbreviations and other confusing requirements. The automated system is, therefore, divided logically into two subsystems: the clerical system and the data processing system. Only very limited familiarity with the computer system is required on the part of library personnel, and the two systems meet only through the interface required to convert data to machine readable form. It should also be understood in considering the approach outlined in this paper that the system is designed as a compromise between the ideal environment of a random store of data accessible on a real-time basis and the actual fact of presently available equipment. While the eventual goal, based on expected future hardware delivery is the more sophisticated operation just mentioned, the system is planned, for the present, to provide the most efficient operation possible utilizing the capabilities of a computer located at a physically remote facility with only limited processing time available for library operations.

File Organization. Each item of data must be uniquely identifiable and individually retrievable. The master tape file is composed of a master unit record for each entry. Desired items of data will be grouped in various printed outputs for use as working files by Serials Department personnel. The master tape record for each entry is grouped logically into several sections: publishing and frequency data, financial and ordering data, bound holdings data, and bibliographic data, for example. In addition, some master records will contain only one section, cross reference data, to the exclusion of all others. This method of processing also requires that each copy of a serial subscription be represented by a separate master record with the appropriate logical sections.

Each master record is identified by a unique six-digit serial identification number (SIN). This number is assigned to the master records so that numerical arrangement by SIN will produce an alphabetical arrangement of entries in the Serials Record. In addition to its use as an alphabetical sequence number, the SIN serves as the unique sequential processing number required for a magnetic tape file application. The SIN was initially assigned at intervals of 100 to allow future insertion of new entries in proper sequence. Whenever an individual interval becomes full, the computer renumbers the entire master file, re-establishing the required interval.

Individual data elements are stored and processed by the computer as both fixed and variable length alphanumeric character strings. A system reserved symbol, the dollar sign, is used to indicate the length of the variable length strings and is the only character not available for use within a string. The character strings are uniquely identified by the SIN and a keyword which has a high mnemonic value for use by Serials personnel.

The keyword is a ten-digit maximum English language word used to identify individual character strings. There is also a special class of keywords that do not have character strings attached but are used to set flags for control purposes. Allowable keywords (see Figure 8) comprise a Serials Record "dictionary" which is open-ended. Addition of new data elements, for example, may be provided by simple modification of the program's keyword table. Keywords for each logical record section of the Master Record are additionally divided into two groups: required keywords and optional keywords. Extensive editing is built into the program and all required keywords must be present, for example, for the particular record section to be accepted. Presence or absence of optional keywords is governed by the nature of the individual entry.

An additional open-ended feature is the ability to introduce an English language "note" into the master file whenever necessary. The keyword XNOTE is used, where "X" represents the first character of the name of the logical record section in which the note is to be included. This feature is the fail-safe measure; if all else fails, hopefully the file can contain a short note to the effect that manual processing is the preferred course of action. The master records are maintained in sorted keyword within SIN order. Bound and unbound holdings record sections are additionally sorted, however, by the major sequence number for the particular title as included in the bound character string. Holdings data thus also require interpretation of the character string for unique identification.

Input Requirements. Input to the computer system is presently in the form of punched cards. The serial identification number (SIN) occupies cc 1-6. The keyword (KYWD) occupies cc 15-28. Card columns 13 and 14 are used for sequence numbering (SEQ) of input

cards when variable length character strings extend beyond one card. Card columns 25-80 contain the character string (DATA). Card columns 7-12 contain a six character maximum operation codeword (OP). This codeword instructs the computer program to perform one of several operations such as ADD, DELETE, or UPDATE.

Clerical procedures require data collection forms similar to that shown in Figure 9. Data are recorded in a manner familiar to clerical personnel whenever feasible. Months, for instance, are entered as common three letter abbreviations and converted to the numeric representation required for calculation. Figure 10 is a sample listing of input cards.

As soon as practicable, a paper tape typewriter will be used in the Serials Department instead of the hand written data collection form. The typed copy will serve to verify data to be entered in the computer file and will serve as the necessary audit trail until the master file is updated. The punched paper tape produced simultaneously will then be used as direct input instead of punched cards. Availability of a keypunch, however, will always be essential for processing punched card arrival files for current receipts.

Output. The output can be divided into two categories: punched and printed. Printed outputs include both the working files for the Serials Department as well as the various listings for public service areas. The working files will be produced monthly and will consist of four physical files: (1) subscription information, (2) bound holdings and bibliographic data, (3) current and unbound items, and (4) a want list. The bound holdings file, for example, will describe each bound volume individually, indicating accession number, inclusive dates, and published volumes included. The computer will produce an item by item shelflist record for inventory control by the Serials Department (see Figure 11) while any holdings information produced for public service areas will probably be printed as inclusive holdings in the form of a union list. In addition to regularly scheduled output, the number of special reports and statistical listings which can be easily produced from the machine records is limited only by the ingenuity of the programmer and the amount of storage space in the Serials Department.

Direct subscription renewals will be initiated by a monthly listing indicating those titles requiring attention. The annual list of domestic serials required for bidding purposes in selection of a state contract jobber will also be printed on continuous multilith masters and reproduced in the required number of copies. The possibility of completely automatic subscription renewal with computer printed purchase orders seems remote, however, considering the complexity of existing state regulations.

The punched card output is used for inventory control and check-in of expected arrivals. At the beginning of each month, one card is

produced for each item expected during the month. Prepunched into each card will be the serial identification number, the operation code-word UPDATE, the keyword UNIT, and a description of the serial issue. Information printed on the arrival card for use by clerical personnel will include this date, plus all other information required to check-in and distribute the issue. This information may also include an optional forty character note, such as "ROUTE" or "CHECK LOOSE INSERTS." As the card is produced, the SIN punched into the card is checked to insure that the correct information is being printed on the card.

Arrival cards are maintained in two separate card files: cards for expected arrivals and cards for received arrivals. Clerical procedure involves matching serial issues and marking them for distribution, as well as the transfer of the corresponding arrival card to the received file. Library files thus show serial holdings as issues expected or issues received in the card files, as issues recorded in the unbound holdings record, as bound volumes, or as missing items appearing in the want list.

One arrival card will be produced for every anticipated individual item. An index, for example, will have a separate arrival card if it arrives separately, or it will be indicated on the arrival card for the issue in which it is bound. Predictable arrivals of supplements will be indicated by separate arrival cards. Each copy of a serial subscription would also require separate arrival cards. A serial subscription for two copies with index separate, for example, would require four arrival cards for the last issue (see Figure 12).

All active subscriptions will be coded as regular or irregular arrivals. A regular subscription is by definition any predictable pattern of publication for which arrival cards can be produced. All other irregular serials with completely unpredictable frequency will be indicated by a printed list. When received, these unexpected items will require manual keypunching of an arrival card to update the master file. Claiming of delinquent arrivals will be aided by inspection of the arrival cards remaining in the file at the end of each month. The decision to claim missing material, however, must be based on the knowledge and judgment of an experienced staff member. "Claimed" cards are flagged and returned to the expected arrivals file.

A second punched output planned is a bindery "tickler" card. This card will be produced in a preassigned month so that the binding load is staggered evenly over a twelve-month period. The bindery clerk will receive a card at the beginning of each month for each title to be bound. The card will list receipts as recorded by the computer. These cards constitute a "tickler" file only; if visual inspection indicates missing issues or insufficient bulk for binding, the card will be refiled until the missing material is received or until the next bindery card for the title is produced.

Serial issues are normally bound into physical volumes of a size to "fit the hand." The decision to bind a given number of issues into a physical volume must be based on the judgment of the bindery clerk. In addition, physical size of individual issues may change drastically and unexpectedly, requiring an exception to the normal routine frequency of binding. The backlog of cards in the bindery file will also aid the clerk in estimating the backlog of unbound material in the collection.

When fully implemented, the computer system will automatically compile monthly operating statistics in addition to routine maintenance of the master file just described. The system is being programmed for an IBM 7094-1401 combination, which includes a model 1404 printer off-line. All non-1401 programs are being written in COBOL whenever possible to insure compatability both for the expected expansion to IBM system 360 equipment, and for the different computer configurations at Rice University and the University of Houston.

The computer facility is made available to other University departments by the Data Processing Center at a reduced rate, and the Center operates as a service facility for University operations as well as serving as an educational facility. For present purposes the Library's need is paramount, and no attempt is being made to present a solid economic justification for the system which is being developed.

Present Accomplishments. In preparation for the approaching conversion of the Serials Record, a Serials Data Processing Clerk was employed as a full-time staff member in the fall of 1965. This young lady, working under the direction of the Serials Librarian, acts as the interface between the sometimes unintelligible serials record files and the keypunch operators by re-ordering and organizing the information on a form suitable for keypunching.

Any data input to a machine records system must be complete and accurate, particularly bound and unbound holdings information used for inventory control. To achieve this lofty goal the Library also employed last fall a crew of full-time clerks to begin a physical inventory of holdings to verify the existing but unreliable records. Holdings information will be converted to machine readable form only after being verified by the inventory crew. It is expected that this project will last at least another year.

The conversion of other Serials Record data to machine readable form was begun in February 1966. Initially, a serial identification number, location, call number, title, and the control keywords, SERIAL or PERIODICAL and ACTIVE or INACTIVE were coded for each of the approximately 7000 entries in the serials holding file and transferred to a magnetic tape file. All cross references were also keypunched and added to the file. In addition, the necessary procedure was established in the Serials Department to insure that title changes and new subscriptions would henceforth be keypunched so

that an updated file could be maintained. All keypunching is being done as much as possible by the regular keypunch operator during the day, who is also responsible for keypunching for the Circulation Department as well as all other library work. A second keypunch operator is employed part-time on a second-shift basis to cope with the volume of data waiting to be keypunched.

Two listings were immediately produced from the data. The first was what might be called a Union List of Serial Titles for Texas A & M University Library. No holdings information whatsoever is reflected, but all cross references are included. This listing, which up to now has been printed only once on Multilith masters and reproduced for campus distribution, will be reprinted only as the need dictates.

The second listing includes only current periodical subscriptions by current title; no cross references are provided. This listing contains the approximately 3500 titles which are available in the current periodical reading rooms in both main and branch libraries, and in the main library Reference Department. It is available in several copies in public service areas in both main and branch libraries and is reprinted presently on a twice monthly basis (see Figure 13).

A little imagination will suggest numerous other uses for the presently limited amount of data now available in machine readable form. One can get simple statistical counts by selected characteristics or listings arranged by call number or location, or combinations of the above. The question is, Do we really want this particular breakdown at this point? There is a temptation to bury the operation in mounds of needless printouts of unwanted information.

The next step is the complete conversion of financial and subscription information, as well as beginning in earnest the conversion of the verified holdings records for the inventoried portion of the collection. The Data Processing Clerk records the ordering and subscription data on mimeographed data collection forms, which are batched and sent to the keypunch operator. Meanwhile, holdings information is keypunched directly from shelflist cards in their catalog trays. As the ordering and subscription data are converted, the computer will immediately begin to flag subscription renewals, and the periodicals list required by the state for bidding purposes will be produced in August 1966 from the information in the magnetic tape file. Our approach to conversion and to the associated experimental programming and debugging is that each mass conversion of a segment of the Serials Record will justify its own existence by immediate utility. Long before complete holdings information is available in machine readable form, the automated serials system will be performing nearly all of the other functions outlined previously.

Conclusion

The progress made in applying data processing techniques to library operations in less than two years is frankly astonishing. One system operates beyond expectations, and a second major operation should be successfully automated within the next year or two. Several lessons have been learned in the course of the last two years. There is no such thing as a completely accurate record where serials are concerned. Attempting to convert data for a computer operation tends to spotlight any weakness in accurate recording of information and, for that matter, in the procedures which supposedly insure that this accuracy is maintained.

There is an old axiom that a computer can best be considered as exceedingly stupid, but it can be letter-perfect where repetitive filing and record maintenance are concerned. We are convinced that perfection is a gift with which mere humans are not yet ready to cope; the perfect error-free input called for cannot be presented, and consequently we spend quite a bit of time correcting mistakes that we created ourselves only the run before. We have also discovered the great need for rethinking what must be done. It is easy to mechanize existing processes, without reviewing the whys of the operation or, in many cases, its usefulness in the first place.

Finally, we believe that our progress can be substantially increased if the library staff is given instruction in basic punched card and computer concepts and then kept abreast of the immediate as well as overall objectives. To accomplish this, a continuing program of lecture and demonstration meetings has been established in which the professional librarians and most of the clerical employees participate. They are encouraged to read new publications of potential interest. In the past, representatives of other libraries have also been invited to attend the two major staff seminar meetings.

REFERENCE

1. Stewart, Bruce W. A Computerized Serials Record System for the Texas A & M University Library. College Station, Texas, Texas A & M University Library, 1965.

DATA PROCESSING CENTER

DPC CONTROL NUMBER: _____
PROGRAMMER: _____
PAGE ____ OF ____
CARD TYPE: _____

LIBRARY MASTER BOOK CARD

1
2
3
4
5
6
7
8
9

(7) (12) (9) (17)

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

1
2
3
4
5
6
7
8
9

Card
Col.

Begins

Shift

2 Item # Field numeric

9 Dewey # Field numeric (col. 12 is always ".")

21 Cutter # Field col. 21 Alpha
22 numeric
23 numeric
24 numeric
25 Alpha
26 numeric
27 "
28 "
29 "

30 "Other" Field numeric

REVISION DATE:

Figure 1
Format and Directions for Key punching the Bookcards

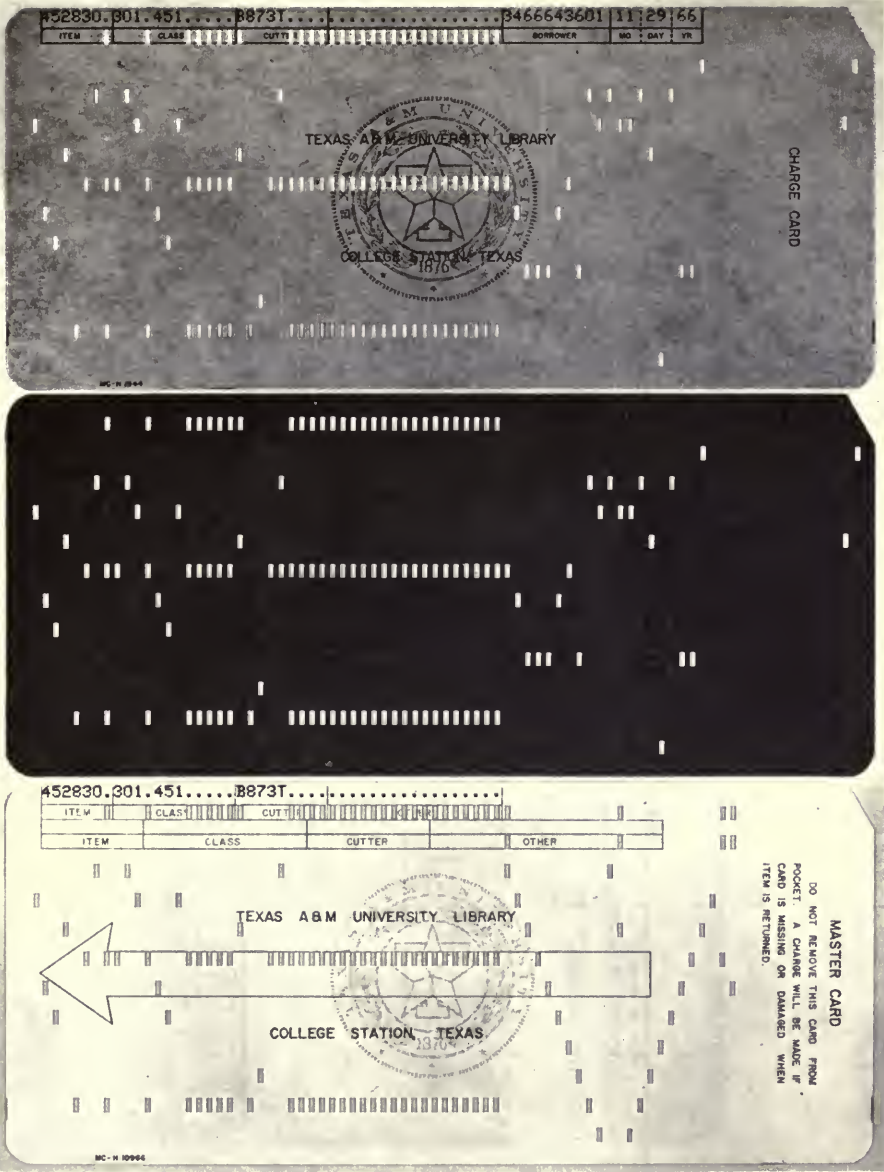


Figure 2
Examples of the Three Punched Cards Used in the Automated Circulation System

DATE OF REQUEST

FOR:

NAME

AUTO CHARGE:

ITEM

MANUAL CHARGE FILE

HOLD UNTIL:

DATE

AUTHOR/VOL

TITLE

C

A

L

L

#

CHARGED TO:

HOLD

CALL IN

000000

111111

222222

333333

444444

555555

666666

777777

888888

999999

71 72 73 74 75 76

Figure 3
Example of Hold Card for Reserving a Book in Circulation

#

BORROWER

#

CALL NO.

Figure 4
Example of Three-part Form Used in Place of the IBM 357 Data Collection System

TEXAS A&M UNIVERSITY MAIN LIBRARY

CALL NUMBER			ITEM---BORROWER---DATE DUE			BOOKS IN CIRCULATION			AS OF 03/12/66			ITEM---BORROWER---DATE DUE		
301.15	C329G	220949	3452583543	03/23/66		301.152	H986B		301.153	1632840000	03/22/66			
301.15	C329G	1960	330814	3396347664	04/10/66	301.152	S291P		309540	1613195000	04/04/66			
301.15	C487P		157762	345629564	05/28/66	301.152	S8731		404472	1612197000	04/09/66			
301.15	C487P	1956	301120	162214000	03/09/66	301.1522	ML335		116613	1604226000	03/23/66			
301.15	C711S		405989	3452583543	03/23/66	301.1523	C743R		314788	1653490000	03/23/66			
301.15	C771T		365737	1638070000	04/04/66	301.1523	H835T		362251	1490157000	04/11/66			
301.15	C771T		365737	1638070000	04/04/66	301.1523	H835T		362251	1490157000	04/11/66			
301.15	C771T		329252	1628070000	04/03/66	301.1523	H831T		294364	1617410000	04/11/66			
301.15	C771T		329252	1628070000	04/03/66	301.1523	H831T		294364	1617410000	04/11/66			
301.15	F332S		266730	1561867000	03/15/66	301.153	B458C		230369	RESERVE	02/28/66			
301.15	F332S		330979	3462093143	03/28/66	301.153	O965		415904	1561867000	03/14/66			
301.15	H274S		257502	RESERVE	11/03/65	301.153	O965		415922	1654154000	03/28/66			
301.15	H332F		215552	1632605000	04/08/66	301.154	C537P		410549	RESERVE	02/16/66			
301.15	H763H		184906	RESERVE	11/03/65	301.154	C347P		305905	1640794000	04/02/66			
301.15	H763H		242822	RESERVE	11/03/65	301.154	L477H		209634	1593939000	04/08/66			
301.15	H763S		412266	RESERVE	02/15/66	301.154	S678P		265965	1630670000	04/07/66			
301.15	K64W	1963	243306	RESERVE	04/11/65	301.15408	G371L		331110	1636860000	03/27/66			
301.15	K64W	1954	243306	RESERVE	04/11/65	301.15408	G371L		331110	1636860000	03/27/66			
301.15	K921		558018	1613164000	TERM	301.155	B364L		343514	1613164000	03/18/66			
301.15	L345RE		235414	RESERVE	11/03/65	301.155	B364L		343514	1613164000	03/18/66			
301.15	L529G		67634	1612197000	04/09/66	301.155	F828H		255189	1650224000	03/25/66			
301.15	L754H	1	241767	3711147623	TERM	301.155	F828H		266794	81N0ERY	01/12/66			
301.15	L754H	2	245475	3711147623	TERM	301.155	G655G		257890	3464323010	TERM			
301.15	M282M	1940	101845	RESERVE	11/03/65	301.155	G655G		364467	3464323010	TERM			
301.15	M479M		102686	RESERVE	11/03/65	301.155	G671L		194006	3462847122	03/20/66			
301.15	R388A		413704	1638070000	04/04/66	301.155	G998S		360303	3462847122	03/20/66			
301.15	R622M		151455	RESERVE	11/04/65	301.155	H151C		300922	3464323010	TERM			
301.15	R622M		151455	RESERVE	11/04/65	301.155	H151C		300922	3464323010	TERM			
301.15	R622M		151457	RESERVE	11/04/65	301.155	H491S		180938	RESERVE	11/03/65			
301.15	R622M		151458	RESERVE	11/04/65	301.155	H737L		404464	3464323010	TERM			
301.15	R622M		151459	81N0ERY	02/09/66	301.155	H945C		229904	RESERVE	02/08/66			
301.15	R622M		151554	RESERVE	11/04/65	301.155	J544		327659	1613164000	03/17/66			
301.15	R622M		261397	RESERVE	11/04/65	301.155	M787P		347878	3464323010	TERM			
301.15	R622M		261397	RESERVE	11/04/65	301.155	M982G		241930	3464323010	TERM			
301.15	R622M		261399	RESERVE	11/04/65	301.155	S554T		267707	81N0ERY	02/21/66			
301.15	R719S		414474	1621310000	04/03/66	301.155	S469L		288862	RESERVE	11/04/65			
301.15	S520	1956	309550	1650933000	03/17/66	301.155	S873M		268493	1650814000	04/10/66			
301.15	S520	1956	327210	1582330000	04/10/66	301.155	U76H		230466	81N0ERY	10/01/65			
301.15	S555S		335922	1582330000	04/10/66	301.155	W754L		208057	3464323010	TERM			
301.15	S6772K3		300723	81N0ERY	02/09/66	301.1574	H741E		322141	1650898000	03/16/66			
301.15	S792P		266274	1622694000	03/24/66	301.1584	H462H		300724	3464323010	TERM			
301.15	S792P		299383	1613195000	04/08/66	301.1583	S966W		357805	3464323010	TERM			
301.15	S792P		330480	1640993000	03/18/66	301.1583	S977C		401520	3464323010	TERM			
301.15	S792P		270481	1432766000	03/18/66	301.1583	T789C		247263	3464323010	TERM			
301.15	S792P		276182	1631652000	03/28/66	301.1584	H741E		297571	81N0ERY	01/12/66			
301.15	S792P		289331	RESERVE	02/28/66	301.1584	H741E		322142	81N0ERY	02/07/66			
301.15	S792P		411941	3456329564	05/28/66	301.1584	H741E		394016	5457584057	03/19/66			
301.15	Y73S		83585	1632605000	04/08/66	301.1584	R643L		191815	3464323010	TERM			
301.15	Z22C		405953	1622435000	04/10/66	301.1584	R643P		264624	3464323010	TERM			
301.1508	F771M		314315	RESERVE	02/04/66	301.1584	R659C		193535	3464323010	TERM			
301.152	C678A		406128	1613195000	04/04/66	301.18	K185S		403726	3389186733	03/27/66			
301.152	H986B		301151	3452126572	03/23/66	301.183	E85C		415233	RESERVE	01/04/66			
301.152	H986B		301152	1602109000	03/10/66	301.183	E85C		415234	RESERVE	01/04/66			

1962

Figure 5
Sample Page of List of Books in Circulation

TEXAS A-M UNIVERSITY LIBRARY							AS OF 03/10/66
DAILY AUTOMATED CIRCULATION STATISTICS							
	UNIVERSITY STUDENTS	UNIVERSITY FACULTY/STAFF	FACULTY/STAFF FAMILY	LOCAL AREA USERS	OTHER USERS	TOTAL	
CHARGES PROCESSED	292	77	0	44	0	413	
RETURNS PROCESSED	305	44	0	25	1	375	
BOOKS IN CIRCULATION	4,725	2,048	123	433	64		
					**** TOTAL CIRCULATION	7,393 ****	
(PLUS 5,385 LOCATION CHARGES)							
OVERDUE BOOKS NOT RETO							
- FIRST NOTICE SENT	1	12	0	0	0	13	
- SECOND NOTICE SENT	3	1	0	0	0	4	
- OVERDUE 30 DAYS PLUS	50	1	3	0	0	115	

CIRC SYSTEM DIAGNOSTIC MESSAGES.												AS OF 03/10/66	
THE FOLLOWING COMMENTS INDICATE PROBLEMS ENCOUNTERED WHILE UPDATING THE CHARGE FILE AND PREPARING OVERDUE AND FINE NOTICES.													
SEVERITY LEVEL PRECEDES EACH COMMENT. (WARNING - THE RECORD WAS PROCESSED AS INDICATED, BUT CORRECTION MAY BE NECESSARY.)													
(DISASTER - MACHINE UNABLE TO COPE WITH SITUATION.)													
ERROR	UNRECOGNIZABLE DUE DATE IN RETURN CARD --- CARD IGNORED. CORRECT AND RESUBMIT. (ITEM#23475, CALL NO 674).....B915LD.....BDRROWER 1653615000* TERM FLAG=0* DATE DUE=05/03/61*)												
ERROR	UNRECOGNIZABLE DUE DATE IN RETURN CARD --- CARD IGNORED. CORRECT AND RESUBMIT. (ITEM#11402, CALL NO 983).....F2837.....BDRROWER 3466540736* TERM FLAG=0* DATE DUE=03/01/66*)												
ERROR	CARD IGNORED. CORRECT AND RESUBMIT. BDRROWER 2621310006* TERM FLAG=0*												

Figure 6
Sample Page of Daily Circulation Statistics and Error Messages

3490

OUR RECORDS INDICATE THAT THE FOLLOWING BOOK IS OVERDUE AT TEXAS A&M UNIVERSITY MAIN LIBRARY. PLEASE PRESENT THIS NOTICE WHEN RETURNING BOOK.

CALL NUMBER
832.....G599FH.....

ITEM NO	DATE DUE	6462846350
247710	12/01/65	MISS JEAN M MARTIN 210 FAIRWAY BRYAN TEXAS

(NOT RETURNED AS OF 12/08/65)

OUR RECORDS INDICATE THAT THE FOLLOWING BOOK WAS RETURNED 9 DAYS OVERDUE AT THE TEXAS A&M UNIVERSITY MAIN LIBRARY. YOUR DELINQUENT FINE IS \$2.25.

CALL NUMBER
949.5.....F688.....

ITEM NO	DATE DUE	DATE RETO	638202
21177.	11/29/65	12/08/65	MR RICHARD A JOHNSON 2801 N HWY 6 BRYAN TEXAS

PLEASE CONTACT THE MAIN LIBRARY CIRCULATION DESK IN ORDER THAT WE MAY RESOLVE THIS MATTER TO THE SATISFACTION OF ALL CONCERNED. UNPAID FINES ARE BILLED THROUGH THE FISCAL OFFICE FOR COLLECTION.

OUR RECORDS INDICATE THAT THE FOLLOWING BOOK WAS RETURNED 9 DAYS OVERDUE AT THE TEXAS A&M UNIVERSITY MAIN LIBRARY. YOUR DELINQUENT FINE IS \$2.25.

CALL NUMBER
949.5.....F688.....

ITEM NO	DATE DUE	DATE RETO	638202
21177.	11/29/65	12/08/65	MR RICHARD A JOHNSON 2801 N HWY 6 BRYAN TEXAS

PLEASE CONTACT THE MAIN LIBRARY CIRCULATION DESK IN ORDER THAT WE MAY RESOLVE THIS MATTER TO THE SATISFACTION OF ALL CONCERNED. UNPAID FINES ARE BILLED THROUGH THE FISCAL OFFICE FOR COLLECTION.

... FILE COPY ...

Figure 7
Sample Page of Overdue Notices Prepared by the Computer

DATA PROCESSING CENTER

OPC CONTROL NUMBER: _____
PROGRAMMER: _____
PAGE ____ OF ____
CARD TYPE: _____
SERIALS RECORD INPUT

OP Codewords:

NEWSIN
ADD
DELETE
CHANGE
UPDATE
BLANK

SIN	OP	SEQ	KYWD	DATA
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				
0				
1				
2				
3				
4				
5				
6				
7				
8				
9				</

DATA FORM

SERIALS RECORD

Keypunched _____ Date _____

SIN 224300 Page 5 Of 7

SIN
224300

OP
ADDL

1 6 7 12

NEW ENTRY -ORDER

Complete For Original Copy Only

Required

-CLASS PERIODICAL

-TYPE RENEWAL

-SOURCE B.I.O. LIST

-LANGUAGE ENGLISH

-PUBLISHED UNITED STATES

-COPIES 2

-INDEXED

-MEMBER AMERICAN MANAGEMENT ASSOCIATION

-DOCUMENT

-FUND 18130-5003

fund order

-PAYOK 01105165

-RENEW 408 month

-PRICE \$ 25.00

-ADSBILLS \$

-PREVIOUS \$ 25.00

-ORDERED 091031044-01015, V54

date material and/or period covered

-BIBNOTE

-NOTE Attached

Figure 9
Form for Preparing Serials Record Input

107840NEWSIN		
107840ADD	XREFER	AMERICAN MANAGEMENT REVIEW. SEE MANAGEMENT REVIEW.\$
108350NEWSIN		
108350A00	1XREFER	ASSOCIATION OF CORPORATION TRAINING. BULLETIN. SEE
108350ADD	2XREFER	MANAGEMENT REVIEW.\$
238350NEWSIN		
238350A00	XREFER	CORPORATION TRAINING. SEE MANAGEMENT REVIEW.\$
784350NEWSIN		
784350A00	1XREFER	NATIONAL ASSOCIATION OF CORPORATION SCHOOLS. BULLETIN.
784350ADD	2XREFER	SEE MANAGEMENT REVIEW.\$
865450NEWSIN		
865450ADD	XREFER	PERSONNEL ADMINISTRATION. SEE MANAGEMENT REVIEW.\$
778300NEWSIN		
778300A00	.ACTIVE	\$
778300ADD	.AODED	JAN65\$
778300A00	.BEGIN	JAN\$
778300ADD	.BINO	001000000000\$
778300A00	.CURRENT	V54N1Y1965\$
778300A00	.CUTTER	M266\$
778300A00	.DEWEY	650.5\$
778300ADD	.ENTRY	MANAGEMENT REVIEW.\$
778300A00	.FREQUENCY	REG111111111111\$
778300ADD	.KIND	ORIGINAL\$
778300ADD	.LOCATE	H-SS\$
778300ADD	.NUM/VOL	12\$
778300ADD	.NUM/YR	12\$
778300ADD	.TPI	SEPDEC\$
778300ADD	.CLASS	PERIODICAL\$
778300ADD	.COPIES	2\$
778300ADD	.FUND	14120-5003\$
778300A00	.LANGUAGE	ENGLISH\$
778300ADD	.MEMBER	AMERICAN MANAGEMENT ASSOCIATIONS\$
778300A00	.ORDERED	09/03/64J-01965,V54\$
778300ADD	.PAYOK	01/05/65\$
778300A00	.PREVIOUS	25.00\$
778300A00	.PRICE	25.00\$
778300A00	.PUBLISHED	UNITED STATES\$
778300ADD	.RENEW	AUG\$
778300ADD	.SOURCE	8ID LIST\$
778300A00	.TYPE	RENEWAL\$
778300A00	80UND	25831V1,1914\$
778300A00	80UND	25832V2,1915\$
778300A00	80UND	25833V3,1916\$
778300A00	80UND	25834V4,1917\$
778300A00	80UND	154352V5,1918\$
778300A00	80UND	154353V6,1919\$
778300A00	80UND	154230V7,1920\$
778300A00	80UND	154231V8,1921\$
778300A00	80UND	375602V9-12,1922-23\$
778300A00	80UND	243104V13,1924\$
778300A00	80UND	243105V15,1926\$
778300A00	80UND	243106V18,1929\$
778300A00	80UND	243107V20,1931\$
778300A00	80UND	154354V21,1932\$
778300A00	80UND	154355V22,1933\$
778300A00	80UND	154356V23,1934\$

Figure 10
Sample Page of Entries in the Serial Record

BOUND HOLDINGS

778300 MANAG

650.5
M266

MANAGEMENT REVIEW.

HISTORY

V1-7N8, 1914-AUG1920 AS NATIONAL ASSOCIATION OF CORPORATION SCHOOLS.BULLETIN. V7N9-V8, SEP1920-JAN1922 AS NATIONAL ASSOCIATION OF CORPORATION TRAINING.BULLETIN. V9, FEB-APR1922 AS CORPORATION TRAINING. V10-11, MAY1922-MAR1923 AS PERSONNEL ADMINISTRATION. V12-14, APR1923-DEC 1925 AS AMERICAN MANAGEMENT REVIEW.

BOUND

25831	V1, 1914	154360	V29, 1940
25832	V2, 1915	295913	V30, 1941
25833	V3, 1916	295912	V32, 1943
25834	V4, 1917	176760	V33, 1944
154352	V5, 1918	243143	V34, 1945
154353	V6, 1919	295911	V35, 1946
154230	V7, 1920	225311	V36, 1947
154231	V8, 1921	216622	V37, 1948
375602	V9-12, 1922-23	216623	V38, 1949
243104	V13, 1924	243108	V39, 1950
243105	V15, 1926	224965	V40, 1951
243106	V18, 1929	224966	V41, 1952
243107	V20, 1931	243109	V42, 1953
154354	V21, 1932	344483	V43, 1954
154355	V22, 1933	341556	V44, 1955
154356	V23, 1934	341557	V45, 1956
154357	V24, 1935	304755	V46, 1957
176761	V25, 1936	341558	V47, 1958
295888	V26, 1937	341559	V48, 1959
154358	V27, 1938	341560	V49, 1960
154359	V28, 1939	502564	V50, 1961
		586431	V51, 1962

Figure 11
Sample Page of the Bound Holdings File

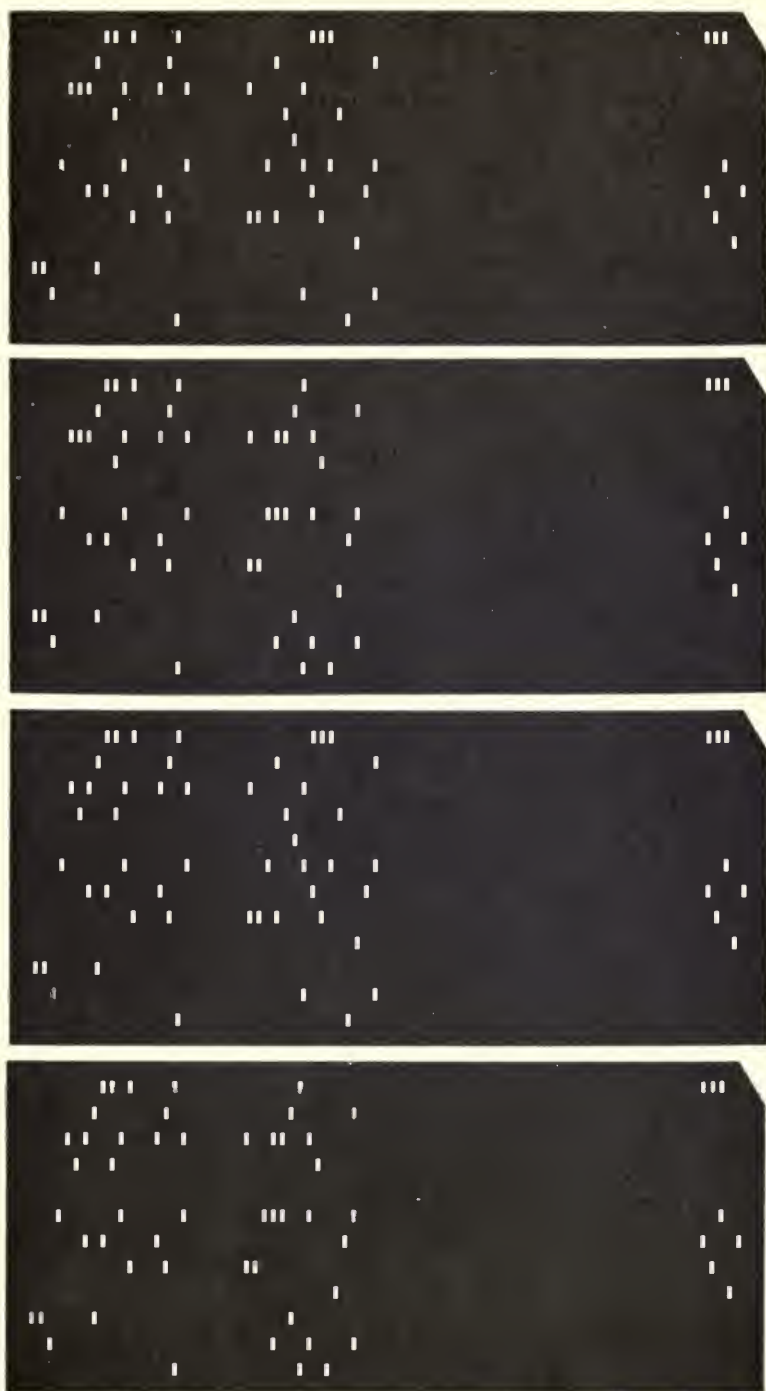


Figure 12
Examples of Arrival Cards for Two Copies of a Serial, with Separate
Volume Indexes

TEXAS A+M UNIVERSITY LIBRARY		FEED
PERIODICAL LIST		
02/14/66		
636.505/H361	FEED AND FARM SUPPLIER	630
636.08505/F295	FEED 8AG	630
636.08505/F2957	FEEDS ILLUSTRATED	630
658.93305/F295	FEEDSTUFFS	650
	FELSMCHANIK UND INGENIEURGEOLOGIE	EL
631.805/F411	FERTILISER NEWS	630
616.6905/F411	FERTILITY AND STERILITY	610
	FERTILITY AND STERILITY	VET
665.05/F421	FETIE-SEIFEN-ANSTRICHMITTEL	660
512.8105/F443	FIBONACCI QUARTERLY	510
677.05/F443	FIBRE AND FABRIC	670
	FICHERO 818LIOGRAFICO HISPANOAMERICANO	REF
799.05/F453	FIELD AND STREAM	700
633.205/C734F	FIELD CROP ABSTRACTS	REF
791.4305/F487	FILMFACTS	REF
332.6305/A532	FINANCIAL ANALYSTS JOURNAL	330
	FINANCIAL EXECUTIVE	8A
332.05/F491	FINANCIAL WORLD	330
	FINANCIAL WORLD	8A
	FINANCING HIGHER EDUCATION	370
	FINISHING ABSTRACTS	EL
551.4606/F511H	FINLAND. HAVSFORSKNINGSINSTITUT. MERENTUTKIMUSLAITOKSEN	
	FIRE ENGINEERING	EL
	FIRE RESEARCH ABSTRACTS AND REVIEWS	EL
812.05/F527	FIRST STAGE - A QUARTERLY OF NEW DRAMA	800
639.205/F537	FISHERIES NEWSLETTER	630
	FLEET OWNER	EL
664.905/F596	DIE FLEISCHWIRTSCHAFT	660
	FLIGHT INTERNATIONAL	EL
	FLIGHT MAGAZINE	EL
580.5/F632	FLORA	580
378.759/F635	FLORIDA. STATE UNIVERSITY, TALLAHASSEE. RESEARCH COUNCIL. FLORIDA STATE UNIVERSITY STUDIES	
506/F636J	FLORIDA ACADEMY OF SCIENCES. QUARTERLY JOURNAL	500
	FLORIDA ENGINEERING SOCIETY. JOURNAL	EL
595.705/F636	FLORIDA ENTOMOLOGIST	590
630.5/F638	FLORIDA GROWER AND RANCHER	630
570.8/F6368	FLORIDA. UNIVERSITY, GAINESVILLE. STATE MUSEUM. BULLETIN	570
639.105/F636	FLORIDA WILDLIFE	630
	FLOW LINE	EL
635.905/F644	FLOWER GROWER	630
	FLUID HANDLING	EL
	FLYING	EL
050/F652	FOCUS (AMERICAN GEOGRAPHICAL SOCIETY)	050
	FOCUS (CONFERENCE BOARD)	8A
574.05/F665	FOLIA BIOLOGICA	570
612.105/F665	FOLIA HAEMATOLOGICA	610
398.05/F666	FOLK-LORE (FOLKLORE SOCIETY, LONDON)	390
398.05/F666F	FOLK-LORE (INDIAN PUBLICATIONS)	390
784.405/F666	FOLKLORE AND FOLK MUSIC ARCHIVIST	700
398.05/F666J	FOLKLORE INSTITUTE. JOURNAL	390
338.105/F636	FOOD AND AGRICULTURAL LEGISLATION	330
338.10611/U57550	FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. FAO AGRICULTURE DEVELOPMENT PAPER	
338.10611/U5755CS	----- FAO COMMODITY BULLETIN SERIES	

Figure 13
Sample Page of the Current Periodicals Subscription List

COMPUTER APPLICATIONS TO BOOK CATALOGS AND LIBRARY SYSTEMS

Donald H. Stromberg

Librarians must look to the future information needs of a country expanding in population, technology and educational requirements. The "information explosion" is placing an additional strain on existing methods of providing information rapidly and economically.

A library seeking to develop a modern information retrieval program has many existing services from which to choose. The problem is to define the program that will best serve the present library users and leave room for flexible action in the future, and then to pick a combination of services that best match these objectives.

Documentation Incorporated (Doc Inc) of Bethesda, Maryland, was founded in 1952 by the late Dr. Mortimer Taube, and has been engaged in developing modern information retrieval systems for government and industry. A key concept that is now emerging is the development of mechanized or computerized data banks. This data bank concept is a plan for organizing a single set of data for producing many products. The traditional library card catalog or data bank, long the key to finding materials in the nation's libraries, today is getting competition from book catalogs. Using data bank techniques to keep a library catalog updated, Doc Inc computer systems generate printouts of the catalog which are used to produce bound books for distribution to library users. In effect, the book catalogs are carrying the traditional card catalog, literally, into the homes and offices of users instead of requiring them to trek to the library to find out if the information they want is available. The computer is used to produce several indexes (such as subject, author, and title) in various formats from a single file of data and is particularly effective if the catalog data bank is standardized.

Despite the wide variety of ways libraries catalog books, most differences are variations of the Library of Congress cataloging system. By using the LC catalog system as a basis or standard for

Donald H. Stromberg is Director of the Publication System Division of Documentation Incorporated, Bethesda, Maryland.

building the data bank, most libraries can be provided with just the index information they need by selecting only those items that apply from the LC data bank. An LC data bank system can produce book catalogs for many libraries from the same data, with provisions for handling special information.

Having introduced the general concept of Documentation Incorporated, the following questions should be answered: (1) Why book catalogs? (2) Why use a computer to produce book catalogs? (3) Why develop a catalog data bank? Each of these questions is inter-related in terms of cost, standards, service, and advancements in technology.

(1) Why Book Catalogs?

There are advantages and disadvantages to both book catalogs and card catalogs. However, it is interesting to note that the card catalog for all practical purposes is a little more than seventy years old and was preceded by book catalogs. In a paper, "Book Catalogs Versus Card Catalogs," by Irwin H. Pizer, the present trend back to the book catalog made possible by recent advances in computer technology is analyzed and advantages and disadvantages are compared.¹ To quote a final statement, "We have gained most of the advantages, have minimized a few of the disadvantages, and, indeed, can look forward like Dr. Pangloss to a computer produced catalog which will be the best possible one, in this best of all possible worlds."

Book catalogs can cut costs by eliminating the purchase or production of cards for multiple locations, the expense of card file maintenance and card cases, and the freeing of prime space now occupied by bulky card catalog cases. They offer more service by providing complete holdings catalogs in bookmobiles; cataloging and order sections; local, school, branch, and cooperating libraries; reference desks; university schools, departments, and residences; and offices of researchers. They can save staff time by eliminating card catalog maintenance activities such as typing cards and headings, sorting, filing, reviewing, and replacing. Book catalogs are as easy to use as a telephone directory—users are generally more familiar with reference books than with card files. Page scanning allows faster location of entries with similar entries and headings, quickly compared at a glance.

Finally, with advancements in computer technology, the quality of the computer final product has been greatly improved. The first advancement was the introduction of both upper and lower case characters directly printable from the computer. The second is the development of photocomposition equipment which is capable of taking

computer generated tapes and composing, with graphic arts quality, the book catalog entries. The quality of the computer produced catalog, which in the past has concerned many librarians, has greatly improved.

(2) Why Use a Computer to Produce Book Catalogs?

Computers are high speed data processors, ideally suited for performing many repetitive clerical tasks, providing that the decisions required can be limited and explained. The scope of the required decisions and the volume of the data to be processed are key factors in deciding what data processing equipment is to be used. It is important to realize that there is a wide range of computers with varying limitations—not the least of which is the craftsman or computer specialist who puts this tool to work. Electronic Accounting Machines (EAM) card processors are quite limited and should not be confused with the memory stored-program computers that utilize tape and disc storage devices as well as punched cards. Card processors have inherent limitations when scanning, compressing, selecting, and sorting data which are required to meet the challenge of the filing rules and variations of catalog information.

Computers are used to produce book catalogs because they offer cost savings, flexibility, consistency, and selected retrieval. The converse of this statement is also true when the tool is applied improperly. An experienced computer specialist would not use this tool solely as a reproducing device. Similarly, a librarian would not send a catalog card to a printing firm to produce one copy. The professional should know his tools and when and how to use them. A computer pays off only when it produces the right information on demand, but what comes out depends not only on what is put in, but on how it is done.

(3) Why Develop A Catalog Data Bank?

The concept of developing a catalog data bank is a key factor in utilizing the computer to produce book catalogs. Cost savings and standardization of catalog information are the major benefits derived from this concept. The Council on Library Resources has granted \$130,000 to the Library of Congress for a pilot program of distributing library catalog data in machine readable form. Under this grant the Library of Congress will also study the value and feasibility of providing this service to libraries on a broad scale and a continuing

basis. Successful completion of this project can reduce even further the cost of computer generated book catalogs, if the primary content of the data bank is standardized to the LC cataloging.

In adopting a cataloging standard such as LC, a library can also experience a substantial reduction in individual cataloging requirements. This is not to say that cataloging will not be required in the future, but since many libraries utilize LC catalog cards in whole or in part, cataloging can be limited to those titles that have not been processed by the Library of Congress.

Professional librarians should have little fear of automation creating unemployment. To quote from the "National Inventory of Library Needs," by Edwin Castagna, concerning the professional staff, "The national shortage in professionally trained staff to meet ALA standards in these libraries would be around 100,000. The data on present staff are for 'professional positions filled' and not professionally trained librarians."² In reviewing the various reports included in the National Inventory of Library Needs, it is apparent from the deficiencies highlighted that there is good reason to limit original cataloging to those items not cataloged by LC if it does not greatly aid the use of the library collection.

Using a Computer to Produce Book Catalogs

This section explains the process used to produce a book catalog from Documentation Incorporated's computer equipment consisting of an IBM 1410 and 1401 tape system. Figure 1 illustrates the general systems flow for computer production of book catalogs. Figures 2-7 are actual samples derived from the Baltimore County Public Library book catalog project.

Each step in the general systems flow is described briefly as follows:

The Catalog Card. The catalog card, which is normally placed in the card drawer, is the primary source of input for computer processing. This is particularly true for those titles that have already been cataloged or are processed by a national organization such as the Library of Congress. When the individual library system performs the cataloging, a standard input form should be used for recording the pertinent data. The use of a standard form with pre-printed identification of each data element to be included in the computer file eliminates the need for additional machine coding and can be used in the ordering process.

To secure a book catalog, a library must provide an identification number for those titles contained in or added to its collection. The identifying number for titles cataloged by LC is the LC card

number and is used by the individual library to call the title from the data bank for the book catalog. Those titles not cataloged by LC require an identification number and processing using the standard input form.

Coding, Editing, and Proofing. The first function in the processing is that of coding and editing the catalog information for machine preparation and is only necessary for those titles not already included in the data bank master file. A typical BCPL entry that has been coded and edited is illustrated in Figure 2. The handwritten marks or "tags" identify each element of catalog information and are used throughout the computer system for selecting, formatting, and generating field descriptions. The standard input form, used by a library to catalog titles that are unique to its collection, should contain the tag identifiers.

Machine Preparation. The machine preparation function converts the written information into machine language. An 80-column Hollerith (IBM) card is the most popular form of computer input. In preparing the catalog entry for the computer system, each tagged item is key punched into a card. If the information for an item exceeds the card limit of 80 columns, a second or third follow-on card is prepared. A special code is punched to indicate that the following character is to be capitalized when required.

The conversion of written information into machine language may also be accomplished effectively using punched paper tape or optical scanning equipment. The state of the art of optical scanning still requires rigid control of the source document format and character style.

Computer Edit and Update. The computer edit and update phase of the processing system performs the file maintenance and is separate from the actual generation of the book catalog. Emphasis is placed on maintaining an accurate data bank file. The five basic steps of this phase follow:

- a. Read the punched cards and create a magnetic tape of the card information.
- b. Sort the magnetic tape into ascending order.
- c. Edit the information to determine if the tags and machine codes are properly encoded.
- d. Update the existing master magnetic tape file with the new information.
- e. Print the proof listing of all changes to the file in upper and lower case.

The IBM 1401 performs steps (a) and (e) and the IBM 1410 performs steps (b)-(d). A new magnetic tape file is created during the updating process and reflects the latest version of the data bank file. This tape is used as input for the next cycle.

A sample of the proof listing from step (e) is illustrated in Figure 3. Errors in coding and machine preparation are corrected on the proof listing, and the file maintenance cycle is repeated. The proof listing contains printed information that is not recorded on the original coded entry. This information is computer generated by identifying the tags and is used to facilitate proofing and proper filing. In comparing subject terms on the proof listing, tag 72001 does not contain the punctuation and capitalization that appear in tag 72002. This is a computer generated sort field and can be changed in the next update cycle if it does not conform to the required filing rules. Sort fields are generated for each major heading and are a very important feature of this system.

Computer Generated Catalogs. Once a corrected master file has been created, the generation of the book catalog becomes a mechanical process performed in the four steps that follow:

- a. Select from the master file the information required for each of the indexes.
- b. Sort the major headings for each index.
- c. Format for printing.
- d. Print the indexes.

It is the decision of the individual library system that determines what indexes are necessary and the contents of each index entry. Figures 4-7 are reproductions of the actual full page IBM 1401 printout for the subject, author, and title indexes requested by the Baltimore County Public Library. Variations of the original entry (see Figure 2) are illustrated in each of the indexes. Examination will also reveal the series statement as a major heading in the title index (see Figure 7).

It is important to reiterate the flexibility that is available in applying computer systems to producing book catalogs. For example, if a library decides that the next publication of the book catalog should not have series statements as a major heading in the title index, there is no need to change any of the entries that have been incorporated in the master file. This is automatically accomplished by instructions to the computer selection program.

Overview Print and Bind. The overview phase of this process is to insure the accuracy of the final product. It is not necessary to cut, paste, or strip the computer print-out prior to making a plate for the printing. In other words, the computer print-out is the camera ready copy. Printing and binding of the final product is the same process as for other types of publications.

LC Entries and Photocomposition. The concept employed to produce book catalogs for the Baltimore County Public Library applied to producing the sample LC entries in Figure 8. The contents of the printed entry in each of the indexes are subject to decisions of

the individual library system. It is, however, more economical to print the full entry in one index and abbreviated versions of the entry in the remaining indexes.

Photocomposition enhances the appearance and reduces the number of pages required in a book catalog. Figure 9 illustrates a sample photocomposed BCPL catalog page. Photocomposition devices, which can be driven from computer generated tapes, offer the flexibility of type styles and graphic arts quality available in typesetting. This process can be justified when the number of pages and copies is substantial, or the publication is semi-permanent such as an annual or biennial cumulative book catalog.

Looking Forward

There are many advantages to automating library systems such as simplifying the ordering and lending processes using computer techniques. Certainly computers can improve reference services by making it possible to provide demand bibliographies. However, most of us are looking to exciting future applications of computers to information processing. Wouldn't it be nice to push a few buttons on the touch-tone telephone that would bring a favorite novel beaming through your television set at home? Impractical? Today, the answer is yes, but look at the developments over the past several decades—television and video tapes, computers linked by telecommunications, touch-tone dialing systems, and last but not least, the catalog card in machine or computer form.

One major step forward is to have a convenient display of the contents of the library collections at the user's disposal and the library collection in a machine form ready to respond to inquiries. The key to matching the user's need for knowledge about the collection and the mechanics of the system is the published catalog. Advancements in technology will make this service economical on a wide basis, but there is a real need for libraries to prepare for the future; publications and educational programs will be required for the users in addition to a conversion of the physical form of the collection into a computerized data bank.

REFERENCES

1. Pizer, Irwin H. "Book Catalogs versus Card Catalogs," Bulletin of the Medical Library Association, 53:238, April 1965.
2. Castagna, Edwin. "National Inventory of Library Needs." In U. S. Office of Education. Library Services Branch, in cooperation with the American Library Association, National Inventory of Library Needs. Chicago, American Library Association, 1965, p. 2.

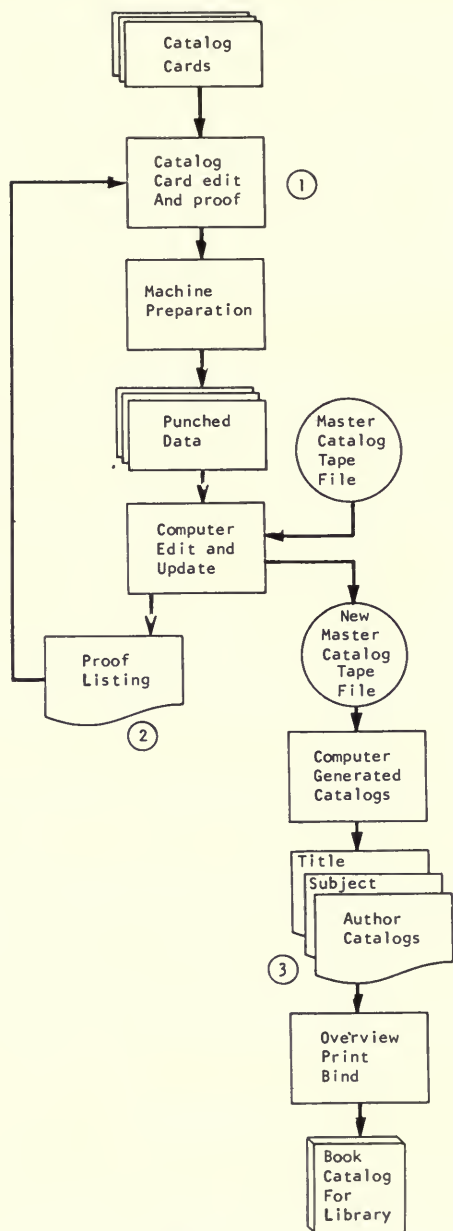


Figure 1
General Systems Flow for Computer Pro-
duction of Book Catalogs

04162	940.542	2400B	66010238
04242	Cole, Hugh Marshall		-48002
	28002	Ardennes: Battle of the Bulge.	1965
	29002	(United States Army in World War II: the European theater of operations)	
		72002	
X	X	Ardennes, Battle of the, 1944-1945	X
		~~Series~~	

Figure 2
A Coded Catalog Card

SUBJECT CATALOG

ART OBJECTS, CHINESE

66011706 595.7 L
APPLICATIONS FOR POSITIONS
Rogers, Nettie How to lead the position you
want 1965.
66011056 658.31 R
AQUARIUMS
Asplund, Herbert R. Trop
66010052
Jenze, A. O. Aquariums
66010756
AQUARIUMS, SALT-WATER
Asplund, Herbert R. Salt-
Rev. ed. 1965.
66010053
AQUATIC SPORTS
Andersen, John Heery Sking on water 3d ed.
1960.
66010023 797 A 1960
ARBITRATION; COMMERCIAL--U.S.
Lazarus, Steven Resolving business disputes
1965.
66010621 369.12 L
ARLEY, FRANCES BURNAY O'
Mellow, Joyce History of Fanny Burnay 1956.
66010659 B A
ARCHAEOLOGY--U.S., EASTERN
Griffin, James Bennett Archeology of eastern
United States 1952.
66010566 917.3 G
ARCHITECTURE
Giddion, Sigfried Architecture you and se
1956.
66010530 720 G
Ruskin, John Seven leaps of architecture 1660.
66011067 720 R
ARCHITECTURE, ANCIENT
Frankfort, Henri Art and architecture of the
ancient orient 1954.
66010491 722 F 1054
ARCHITECTURE, BYZANTINE
Kraushaar, Richard Early Christian and
Byzantine architecture 1965.
66011761 723 K
ARCHITECTURE, COLONIAL
Halsey, Richard Townley Homes of our
ancestors 1925.
66010609 749 N
ARCHITECTURE--DETAILS
Ours, Alison Early ironwork of Charleston
1941.
66010306 730.4 O
ARCHITECTURE, DOMESTIC
Gillies, Mary Davis McCall's book of modern
houses Rev. ed. 1959.
66010541 728 C 1959
Halsey, Richard Townley Homes of our
ancestors 1925.
66010609 749 N
ARCHITECTURE, DOMESTIC--JAPAN
Seika, Kiyosi Contemporary Japanese houses
1964.
66011129 720.95 S
ARCHITECTURE, DOMESTIC--U.S.
Nicholson, Arnold American houses in history
1965.
66010977 917.3 N
ARCHITECTURE--ENGLAND
Clark, Kenneth McKenzie Gothic revival 1962.
66010220 724.3 C 1962
ARCHITECTURE, GOTHIC
Clark, Kenneth McKenzie Gothic revival 1962.
66010220 724.3 C 1962
ARCHITECTURE--IRAN
Pope, Arthur Upham Persian architecture 1965.
66011036 720.956 P
ARCHITECTURE, MODERN--20TH CENTURY
American federation of arts Fountains in
contemporary architecture 1965.
66010016 731.722 A
ARCHITECTURE--RUSSIA
Neslison, George Heerd Art and architecture of
Russia 1964.
66011659 709.47 N
ARCHIVES--PENNSYLVANIA--DIRECTORIES
Pennsylvania. Historical and museum commission

ARDENNES, BATTLE OF THE, 1944-1945
Cole, Hugh Marshall Ardennes: Battle of the
Bulge 1965.
66010238

940.542 C

ARITHMETIC
Eicholz, Robert E. Modern general arithmetics
1965.
66010366 511 E
Slate, Edwin L. Refresher arithmetics with
practical applications 1965.
66011196 510 S
ARITHMETIC--EXAMINATIONS, QUESTIONS, ETC.
Arco publishing company, New York Civil
service arithmetic and vocabulary
CONSULT LIBRARIAN
66010026 351.3 A
ART--ADDRESSSES, ESSAYS, LECTURES
Elliot, Alexander Sight and insight 1959.
66010392 701 E
ART, CHINESE
Boston. Museum of fine arts Charles B Hoyt
collection
CONSULT LIBRARIAN
66011436 709.51 H
Jenyns, Soame Chinese art, the minor arts, II
1965.
66010765 709 J
Willett, William Foundations of Chinese art
from Neolithic pottery to modern architecture
1965.
66011326 709.51 W
ART--COLLECTORS AND COLLECTING
Cooper, Douglas Great family collections 1965.
66010256 708 C
ART, COPTIC
Vesseli, Klaus Coptic art 1965.
66011312 709.32 W
ART--ENGLAND
Thomas, Gertrude Z. Richer than spices 1965.
66011240 914.2 T
ART--HISTORY
Hertli, Michael Rhels Arts 1965.
66010876 709 M
ART INDUSTRIES AND TRADE
Lalor, Allen Haddershoil Handicrafts of New
England 1949.
66010376 745.1 E
ART INDUSTRIES AND TRADE--SWEDEN
Held, Arthur Contemporary Swedish design 1951.
66010605 709.48 N
ART, IRANIAN
Forede, Edith Art of ancient Iran 1965.
66011934 709.35 P
ART, ITALIAN
Cheslail, Andre Golden age of the Renaissance
1965.
66011462 709.45 C
ART, MEDIEVAL
Loomis, Roger Sherman Mirror of Chaucer's world
1965.
66010647 821 L
ART--MESOPOTAMIA
Hargerson, Jean Claude Mesopotamia 1965.
66010870 913.35 M
ART--MEXICO
Crosby-Kisbell, Ingerd Art of ancient Mexico
1954.
66010593 913.72 G
ART, MODERN--20TH CENTURY
Kah, Katherine Woolf Break-up: the core of
modern art 1965.
66011786 709.04 K
ART OBJECTS, CHINESE
Boston. Museum of fine arts Charles B Hoyt
collection

Figure 4
Subject Entry in Subject Index

AMERICAN GOSPEL OF SUCCESS		TITLE CATALOG	
96511499	917.3 B	Bcha191-Nialson, Keat Stortebasher 29 a9. 1994.	
AMERICAN GOPEL OF SUCCESS		96511114	091.1 S 1964
Biechle, Monroe 1995.			
96511096	399.173 R	ANNALS OF SALTIDORE	
AMERICAN HOUSE IN HISTORT		Griffith, Thomas Waters 1824.	
Nicholson, Aroel9 1968.		96511845	H9. Ref. 875.28 G
96510877	917.3 M	ANYTHING BUT	
AMERICAN ILIAO		Neefarson, Zeece 1889.	
Eiseeeh1el, Otto 1947.		96511994	3F
96511359	973.7 E 1947	APPLE OF HIS EYE	
AMERICAN LABOR TODAY		Bohichae9, Garar9 1905.	
Harr, Norbert Louis 1960.		96511079	
96511981	331.88 M	APPLIES BUSINESS LAW	
AMERICAN LITERAYUDE		0 Fish, Hckae 9th a9. 1980.	
Otelinae, Robert Woester 1981.		96511089	347 F 1960
96511155	918.9 B	APPAICING AND BELLING YOUR COINS	
AMERICAN NEGRO, A CHRONOLOGT AND FACT BOOK		Fria9harg, Bohart 4th a9. 1994.	
Bloee, Ivelog J. 1995.		96511902	737.4 F 1994
96511179	Ref. 873.897 B	AQUABIU TECHNIQUE	
AMERICAN NOTEBOOK		Jaena, A. G. 1894. 2v.	
Haeherger, Philip Paul 1960.		96515759	098.7 J
96515618	817.3 M	ARANSIQUE	
AMERICAN POETRY SINCE 1940		Househo19, Boeffrey 1848.	
Stepascher, Stephan 1965.		96511710	
96511269	811.59 B	ARCHAEOLOGY IN THE MOLT LAND	
AMERICAN POETRY		Kaegee, Kathinae Mary 29 a9. 1985.	
Ehneapre1e, Irela 1866.		96511797	913.394 K
96511843	911.95 E	ARCHAEOLOGY OF THE CINEMA	
AMERICAN POLITICS AND THE PARTY SYSTEM		Harck, Kurt U. 1960.	
Boee, Hugh Alsio 3d a9. 1965.		96510990	779.8 B
96510912	32d B	ARCHEOLOGY OF EASTERN UNIVIO STATES	
AMERICAN STEAM VESSELS SERIES, V.1		Griffie, Jaena Beacott 1882.	
Staatoo, Samuel War9 Grant Lohae etoeo vaseelo		96510939	917.3 G
1962.		ASCHIBALS HACLEIBH	
96512521	307 B	Felb, Sigal Loeoe 1860.	
AMERICAN STEAM VESSELS SERIES, V.2		96511859	911 M
Staatoo, Samuel Ward Long Island Sooe9 a9		ARCHITECTURE YOU ANG HE	
Narragoeett Bay etoeo vaseelo 1962.		Bieeloe, Sigfrie9 1909.	
96512822	387 B	96519239	729 G
AMERICAN STEAM VESSELS SERIES, V.3		ARDENNES: BATTLE OF THE BULGE Cole, Hugh Marshall 1965. 66010238 940.542 C	
Staatoo, Samuel War9 Ocoe			
96512824			
AMERICAN STEAM VESSELS SERIES, V.4			
Staatoo, Samuel Ward Mied			
Butter aaeel etoeo vaseelo			
96512023	309.32 B	ARY ANG ARCHITECTURE OF BUDGIA	
AMERICAN STEAM VESSELS SERIES, V.5		Naellitoe, George Haer9 1804.	
Staatoo, Samuel War9 Steamhoete of the river		96511859	709.47 H
Nedeece 1965.		ABT AND ARCHITECTURE OF THE ANCIENT ORIENT	
96512825	397 B	Frankfert, Haeri 1954.	
AMERICANS AND CHINESE		96518481	722 F 1964
Wae, Freaete Long Keeng 1953.		ABY ANG ANDYROL (A C BARHE9)	
96516717	810.1 M	Gchach, Williae 1999.	
AMERICA'S HOMEKAKING BOOK		96511876	S S
Ooe9, Marguerite 1957.		ARY OF ANCIENT IRAN	
96511931	949 S	Para9e, E9ith 1965.	
AMPLIFIED SIOLE		96511934	789.30 P
Silnie 1955.		ART OF ANCIENT MEXICO	
96518162	225.8 B	Groth-Kiabell, lrrgar9 1954.	
ANULETO AND SUPERVISIONS		96518053	913.72 G
Sudge, Ereet AlFred Theapoece Nellie 1939.		ABT OF MEXICAN COOKING	
96518147	390.3 B	Araee, Jan 1965.	
ANARCHISM: A HISTORT OF LISERVARIAN ISEAE ANG		96518081	941.0 A
MOVEMENTS		Loelie Frederic, pae9. 1905.	
Needcoe, George 1862.		96518649	785.59 L
96512117	335.92 M	ART OF THE PUPPET	
ANATONT OF BRITAIN TODAY		Beir9, Bill 1965.	
Saepeeoe, Anthony 1965.		96510006	781.0 B
(Aliee puhlika9 as Anatomy of Britiae)		ARY OF THE NOTAL BALLET	
96511971	814.2 B	Neey, Kaith 1964.	
ANCIENT IRAQ		96518941	793.3 H
Roae, George 1964.		ART, SCIENCE AND SPIRIT OF NURRING	
96511595	513.35 B	Price, Alice Loelee 39 a9. 1965.	
AND IV HAS MORNING		96511540	916.73 P 1965
Borchaeeloe, Pool 1962.		ARTILLES THROUGH THE AGES	
96511434	909.54 S	Staweece, Phillip M. 1995.	
AND THE BAIN NY OBINK		96512031	359.1 B
Hae, Gaele, pae9. 1959.		ARTIST'S LIFE	
96518614		Estare, Angee 1909.	
ANSRE CEMIER		96518454	9 E
Loggiee, Waeoe 1965.		ARTISTS OF THE OLS WEST	
96518122	B C	Eaee9, John Caefield 1965.	
ANIMAL PHYIOLOGY		96518425	769.1 E
Behaldt-Nialson, Keat Stortebasher 1965.		ARTS OF HARKINS	
96511115	891.1 B		

Figure 6
Title Entry in Title Index

TITLE CATALOG		UNIVERSAL DECIMAL CLASSIFICATION	
66611436 TOOL DESIGN Sawedown, Cyril 2d ed. 1967.	155.4 P	Cornes, Robert 1966. H9611478	387.2 C
66611938 TOWNETS OF SPINING Henloggess, Ernest 1933.	621.9 S	TWIM: TWICE THE TROUBLE, TWICE THE FUN Gohaus, Seton Holland 1966. 66915825	156.444 G
66615058 TONTUNE'S NOISE Inman, Robert 1966. H9611728		TWINTED PEOPLE Phillips, Judson Pentecost 1968. H9611828	N
66611276 TOWES OF KILRAVEN Crosse, Cecily 1968. 66611276		TWO ANNAHANS Hartigan, Harriet 1965. 96611872	384.2 H
66611276 TOYLAND Smith, Mark 1966. H9611177		TWO BEATHS OF QUINCA'S WATNEYELL Arends, Jerg 1966. 66611873	
66611277 TMASTIONAL QUILTING Fitzsardolph, Neale 1984.	445.67 F	TWO LIVES OF ESTH WHASTON Griffiths, Grace Kellum 1968. 96611844	B W
66611278 TMAGEST, WIMON AND FOMH Corrigan, Robert Milloughby 1966. H9611253	866.2 C	TWO THOUSAND YEARS OF CALLIGRAPHY Sollinore, Hsueh of art 1966. 96615968	741 B
66611244 TMASTON METRATES Mills, Dunnington, pseud. 1964.	822.44 T	U	
66611078 TMAVEL GUISSE TO THE ORIENT AND THE PACIFIC Velds, Myre 1966.	H	UNCLE JOHN'S ORIGINAL SEAS BOOK Greene, John Wals 2d ed. 1965. 66611448	841.6 S
66611286 TMAVELS OF MARCO POLO, THE VENETIAN Polo, Marco 1936.	914 P 1936	UNCOMMITTES Kumlaten, Kenneth 1965. H96118765	381.2 K
66611932 TREASURY OF LIFE INSURANCE HALEN 1968 Crownin, Hsueh A. 1969.	616 P 1936	UNCOMMITTES AMERICA'S INSURANCES Gerbrach, Carl 1962. 96611819	MMH G
66611933 TRIALS OF THE WOMB Lentz, Michas Harrington Hsueh 1966. H9611931	866.66 C	UNCOMMITTES MACROECONOMICS Hillbrenner, Robert Louis 1965. 96611862	336.1 H
66611931 TRIGONOMETRY Hesse, Dorothy 2d ed. 1969. H9611849	616.6 L	UNCOMMITTES HAGAZINES Wolstein, Roland Edgar 1965. 96612116	861 H
66611849 TRIUMPHANT EMPIRE Gipson, Laurence Henry British empire before the American Revolution CONSULT LUMABIAN 96611849	614 H	UNEXPLORED BONES Hartigan, Arthur Sanford 1968. H9611871	444.44 H
66611849 TRIUMPH OF BIOLOGY Goldstein, Philip 1965. 96611823	942.672 G	UNHOLY ALLIANCE Frande, Gerald 1967. 96618465	327.47 F
66611823 TRIUMPH OF CONQUEST Gould, Cecil 1969. 96611833	974.44 G	UNHOLY UPBORN Savile, Clyde Srio 1967. H9611826	
66611833 TROPICAL FISH BOOK Axirod, Herbert M. 1964. H9611892	766.4 G	UNIONS BEFORE THE SAN Lieberman, Elias Rev. ed. 1966. 56615839	331.626 L
66611892 TROUBLE IN MENDELEY Harrigan, Steven 1969. 96612092		UNITED STATES AND THE HAWAIIAN KINGDOM Tate, Harro 1966. H9611229	327.966 T
66612092 TUMMOTT MUTE Plover, William Charles F. 1966. H9611824		UNITED STATES ARMY IN WORLD WAR II: THE EUROPEAN THEATER OF OPERATIONS Cole, Hugh Marshall Ardennes: Battle of the Bulge 1965. 66010238	
66611824 TUMF MANAGEMENT Hessy, Howard Ertie Rev. ed. 1962. 96610994	942.672 G	940.542 C	
66611824 TUMKET Stewart, George S Stilley 1965. 66611299	915.4 H	PC(2)-MM. HUNJECT MENDONTH: INCOME OF THE ELSENLY POPULATION. U.S. Census, Bureau of the 1963. 96611276	
66611299 TMENTIER: AMERICAN WRITING IN THE POSTMAN DECADE Hoffman, Frederick John 1965. 66611966	915.4 H	U.S. COAST GUARD Capron, Walter C. 1966. 96611471	366.67 C
66611966 TMENTIER-CENTURY HUNOT HUNOT EXPLANATION: SUPPLEMENT II. INTERPRETATIONS, APRIL 1, 1963- DECEMBER 31, 1964 OF SHORT FICTION SINCE 1966 Heller, Warren S. 1965. 96611292	Ref. 915.4 H 1965	U.S. COSE, 1964 ES U.S. Lane, statotes, etc. 1966. 14v. 66612071	Ref. 345 U 1969
66611292 TMENTY DAYS Kunhardt, Dorothy Hsueh 1966. 96611911	8 L	UNITED STATES IN WOMB AFFAIRS, 1964 Savile, Jaleo 1966. (Former author Hsueh, Richard Pote) H9611824	327.73 H 1964
66611911 TMENTY TWENTY-TWO: GOLDEN APPLE OF THE SUN, MEDICINE FOR MELANCHOLY Hsueh, Hsueh 1966. 66611444	MF	U.S. INFANT: QUEEN OF BATTLE Gillman, Richard Joseph 1966. 66612638	366.1 S
66611444 TMENTY OF SAILING SHIPS		UNIVERSAL DECIMAL CLASSIFICATION	

Figure 7
Series Entry in Title Catalog

FISHLOCK, DAVID

Metal colouring. Teddington, R.
Draper, 1962. xii, 393p. illus. (part
col.) 23cm. Includes bibliographies.

TS653.F53 64-35743

OKLAHOMA. STATE UNIVERSITY OF AGRICULTURE
AND APPLIED SCIENCE, STILLWATER. SCHOOL
OF CIVIL ENGINEERING.

Research publication. Stillwater, 19
no.in v. illus. 29cm. Issued by the
school under the university's shorter
form of name: Oklahoma State
University.

TA7.o45 64-35744

UNITED NATIONS. OFFICE OF PUBLIC
INFORMATION

The International Court of Justice.
[4th ed.] New York, United Nations
(1963[41p. illus. 20cm. =United
Nations publications. Sales
no.:63.I.11"

JX1971.6.U47

64-54693

Figure 8
Sample Computer Print of L. C. Entries

AUTHOR CATALOG

HOSKINS, WILLIAM GEORGE

- HENRY FRANCIS DU PONT WINTERHURST
 MELISEL
 Winterthur portfolio one c1964 708 1 H
 0145-12401
 HENRY, O. PSEID D.
 Census reader 1904 706 54 H
 0265-15253
 Heart of the south 1904
 0265-15254
 Ogeon 1909
 0265-15255
 Roads of destiny 1903
 0265-15256
 Seas and rivers 1904
 0265-15257
 Sixty business 1910
 0265-15258
 Tinned lamp 1907
 0265-15259
 Voice of the city 1908
 0265-15260
 HENNER, HARRY WALKER
 Advertising creative communication with consumers
 4th ed. c1964 659 1 H
 0145-12403
 HERBERG, CHARLES GEORGE
 Suburbans in the United States c1916
 0145-12404
 HERBERT, ALAN PATRICK
 Barber M P 1964 271 75 H
 0265-15261
 HERBERT, SIR HENRY
 Dramatic records of Sir Henry Herbert 1917
 0145-12411
 HERLING, JOHN
 Labor actions in America 1964 792 G
 0265-15262
 HERINGTON, CALVIN C.
 Sea and rivers in America c1965 126 1 H
 0265-15263
 HERB, DAN
 Ruffles and sprouts 1964 792 G
 0265-15264
 HERBIE, S. CHARLES JUDSON
 Brains of rats and men c1962 596 H
 0145-12405
 HERBIE, RENNETH W.
 Total disability provisions in life insurance contracts c1956 368 3 H
 0145-12406
 HERSEY, JOHN RICHARD
 Campaign for Adano 1948 320 1 H
 0265-15265
 White Lotus 1964
 0265-15266
 HERTZ, LOUIS HEILBRONER
 Combined levels of model railways and roadways
 1964 629 2 H
 0265-15267
 HERZOG, ARTHUR B.
 War plans and strategy 1965 659 133 H
 0265-15268
 HERSHFIELD, WILLIAM BEST
 Lincoln's plan of reconstruction c1960 973 7 H
 0145-12408
 HERBSON, JACQUES
 Daily life of the Etruscans c1964 913 37 H
 0145-12409
 HERSEVELD, HELEN
 100 best places to live when you retire 1964 913 3 H
 0265-15270
 HERRATT, ALEXANDER
 Historical account of the rise and progress of the colonies of South Carolina and Georgia 1779 2c
 0145-12410
 HEVEL, CARL
 Appraising executive performance c1958 658 3 H
 0145-12401
 HEYER, ANNA HARRIET
 See American library association c1962 658 3 H
 HEYWOOD, JOHN
 Radio astronomy and how to build your own telescope c1963 523 H
 0145-12411
 HIBBERT, CHRISTOPHER
 Court at Windsor 1964 920 H
 0145-12412
 HICKIN, RAL
 Union vs Dr Mudd c1964 8 36
 0145-12413
 HILDRETH, GERTLIE HOWELL
 Teaching reading: a guide to basic principles and modern practices c1958 372 4 H
 0145-12414
 HILL, JON DAN
 Menace man in peace and war; a history of the National Guard c1964 355 3 H
 0145-12415
 HILL, NAPOLEON
 Think and grow rich c1937 174 H
 0145-12416
 HILL, RALPH NADING
 Mid-century drive being an account of the first auto trip across the U.S.A. San Francisco to New York, 1902, or, fifty-three days on a Wilson motor carriage c1964 913 3 H
 0145-12417
 HILL, WILLIAM H.
 A memoir Bradburn, his life and work (1644-1737) c1962 787 12 H
 0145-12418
 HILLARY, EDMUND
 Schoolhouse in the clouds 1964 915 496 H
 0265-15272
 HILLARY, LOUISE
 Keep calm if you can c1964 796 54 H
 0145-12419
 HILLICOURT, WILLIAM
 Roder Powell: the two lives of a hero c1964 R R
 0145-12420
 HILLWAY, TYRIS
 Education in a American society, an introduction to the study of education c1961 370 1 H
 0145-12421
 MINDE, THOMAS (PSEID D.)
 See City, Taiwan
 MINNAN, CHARLTON
 Pressing and proof-reading of the first folio of Shakespeare c1963 2c
 0145-12422
 HIRSHMAN, ALBERT O.
 Journeys toward progress: studies of economic policy-making in Latin America c1963 338 98 H
 0145-12423
 HISTORY OF CIVILIZATION
 History of civilization, ed by C H King 2nd ed c1964 2c
 0145-12424
 HISTORY OF ROME AND THE ROMANS, FROM
 ROMULUS TO JOHN XXIII
 History of Rome and the Romans, from Romulus to John XXIII General director, Robert Laffont c1960 965 832 H
 0145-12425
 HITLER, ADOLF
 Mein Kampf 1964 960 54 H
 0265-15273
 HORTSANG, CARL
 A master chess director c1961 784 H
 0145-12426
 HOAG, EDWIN
 American houses, colonial, classic and contemporary c1961 728 H
 0145-12427
 HOBART, ALICE TISSAULE NOURSE
 Child recs 1964
 0265-15274
 HOBBS, LEONARD TIBELMAN
 Liberalism 1964 320 1 H
 0265-15275
 HOBSON, LAI BA KEANE TSEH
 First papers c1964
 0145-12428
 HODDER, ANDRE
 Jazz, its evolution and essence c1956 781 57 H
 0145-12429
 HODGSON, JOHN H.
 Earthquakes and earth structure c1964 551 2 H
 0145-12430
 HODGSON, RICHARD S.
 Dartford direct mail and mail order handbook 1964 659 133 H
 0265-15276
 HODGES, GERNARD E.
 Introduction to Boolean algebra and logic design: a program for self-instruction c1964 512 89 H
 0145-12431
 HOFER, ERNST
 Actor reviews c1957 910 8 H
 0145-12432
 HOFER, ARTHUR S. LEVILLANT
 Writing of fiction c1954 808 3 H
 0145-12433
 HOFFMAN, MARTIN LEON
 Review of child development research 1964 1 16 7 H
 0265-15277
 HOFFMAN, NICHOLAS VON
 Manuscript notebook 1964 323 H
 0265-15278
 HOFMEIER, JOHANNES
 A art of teaching Christian doctrine 2nd rev and ext ed c1962 238 2 H
 0145-12434
 HOFMANN, JULES
 History of Zion Church of the city of Baltimore, 1755-1907 1905 166 284 1 H
 0145-12435
 HOFSTADTER, RICHARD
 Great issues in American history 1958 2c
 0265-15279
 United States, history of a republic c1957 973 H
 0145-12436
 HOGAN, ROBERT GOODE
 A ruler Mithra c1964 812 H
 0145-12437
 HOGARTH, PAUL
 Crayon pencil drawing c1964 741 H
 0145-12438
 HOGG, GARRY
 Hammer and song 1964 662 H
 0265-15280
 HOGNER, DOROTHY CHILDS
 Gardening and cooking on terrace and patio c1964 635 H
 0145-12439
 HOFENBERG, JOHN
 Foreign correspondents: the great reporters and their times c1964 070 431 H
 0145-12440
 HOLCOMB, RICHARD LORAN
 Police and the public c1954 352 2 H
 0145-12441
 HOLDS, JOHN L.
 Young man's guide to camping c1962 786 54 H
 0145-12442
 HOLDER, ROBERT
 Complete guide to successful school publications 1964 371 82 H
 0265-15281
 HOLLE, CHRISTINA
 Christmas and its customs, a brief study c1957 394 H
 0145-12443
 HOLLAND, LYNWOOD MATHIS
 Peter M B Young, the Warwick of the South c1964 R Y
 0145-12444
 HOLLANDER, LEE M.
 Sage of the Keweenaw c1955 839 8 J
 0145-12445
 HOLLEY, HENRY BRINTON
 Boying army: financial procurement for the Army and Forces 1964 (U S Army in World War II Special studies) 358 4 H
 0265-15282
 HOLLES, CHRISTOPHER
 Papey, an illustrated history from St Peter to Paul VI c1964 262 13 H
 0145-12446
 HOLMAN, MARY
 Police officer and the child c1962 364 44 H
 0145-12447
 HOLMES, OLIVER WENDIE J.
 Holmes's station letters: correspondence of Mr Justice Holmes and Lewis Luntz 1903-1915 c1964 R H
 0145-12448
 HOLST, THOMAS
 A.R. of music 1961 781 H
 0265-15283
 HOLY, JOHN A. ALMOND J.
 How children feel c1964 372 H
 0145-12449
 HOLY, P. J. M.
 A Columbia history and literature 1964 918 81 H
 0265-15284
 HOLY, S. H. HOLY
 Logical and logical logic c1965 781 H
 0145-12450
 HOLTON, DANIEL CLARENE
 Modern, Jacob and Mistic rationalism Rev. ed 1967 299 15 H
 0265-15285
 HOLTON, GRAHAM
 Foundations of modern physical science c1964 530 H
 0145-12451
 HOLTON, HENRY HAN FUR 1964
 Home improving ideas for 1965 1964 47 8 H
 0265-15286
 HOLY IMPROVEMENT IDEAS FOR 1965
 Home improving ideas for 1965 1964 690 8 H
 0265-15287
 HOLMAN, JOHN
 Introduction to the chemistry of carbohydrates 2nd ed 1964 547 78 H
 0265-15288
 HOLMANN, JOHN JOSEPH
 Culture and personality c1954 137 H
 0145-12452
 HOLMAN, D. JOHN
 Life of the U.S. c1964 540 199 H
 0145-12453
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15289
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12454
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12455
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15290
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12456
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12457
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15291
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12458
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12459
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15292
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12460
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12461
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15293
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12462
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12463
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15294
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12464
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12465
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15295
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12466
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12467
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15296
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12468
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12469
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15297
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12470
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12471
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15298
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12472
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12473
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15299
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12474
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12475
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15300
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12476
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12477
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15301
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12478
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12479
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15302
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12480
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12481
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15303
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12482
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12483
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15304
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12484
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12485
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15305
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12486
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12487
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15306
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12488
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12489
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15307
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12490
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12491
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15308
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12492
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12493
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15309
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12494
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12495
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15310
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12496
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12497
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15311
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12498
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12499
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15312
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12500
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12501
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15313
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12502
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12503
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15314
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12504
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12505
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15315
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12506
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12507
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15316
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12508
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12509
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15317
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12510
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12511
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15318
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12512
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12513
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15319
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12514
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12515
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15320
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12516
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12517
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15321
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12518
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12519
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15322
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12520
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12521
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15323
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12522
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12523
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15324
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12524
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12525
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15325
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12526
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12527
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15326
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12528
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12529
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15327
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12530
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12531
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15328
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12532
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12533
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15329
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12534
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12535
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15330
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12536
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12537
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15331
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12538
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12539
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15332
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12540
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12541
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15333
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12542
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12543
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15334
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12544
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12545
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15335
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12546
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12547
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15336
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12548
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12549
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15337
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12550
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12551
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15338
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12552
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12553
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15339
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12554
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12555
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15340
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12556
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12557
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15341
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12558
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12559
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15342
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12560
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12561
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15343
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12562
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12563
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15344
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12564
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12565
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15345
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12566
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821 H
 0145-12567
 HOLMAN, NINA WARNER
 Sea summer 1964 999 743 H
 0265-15346
 HOLMAN, BRIE E.
 Mother and child in modern art 1964 364 94 H
 0145-12568
 HOLMAN, HENRY HOLMAN
 George Orwell c1962 821

LIBRARY APPLICATIONS OF DATA
PROCESSING: AN ENUMERATIVE BIBLIOGRAPHY, 1964-65

Compiled by James Krikelas

The purpose of this bibliography, prepared for the participants of the 1966 Clinic, is to provide a convenient list of references to actual applications of data processing machines for the mechanization of library routines. Items published in 1964 that were included in a similar list distributed in 1965 have been excluded.* In addition to Library Literature and Library Science Abstracts, a new quarterly, Documentation Abstracts (which includes the former "Literature Notes" section of American Documentation), should be useful for current references. For retrospective citations to operational systems the references cited below may be useful. (It should be noted that the scope of each is not identical and some duplication has occurred.)

Helen E. Loftus and Allen Kent. "Automation in the Library—An Annotated Bibliography," American Documentation, 7:110-126, April 1956.

Edward Mack McCormick. "Bibliography of Mechanized Library Processes." In Herbert Goldhor, ed., Proceedings of the 1963 Clinic on Library Applications of Data Processing. Champaign, Ill., University of Illinois Graduate School of Library Science, 1964, pp. 157-176.

Robert S. Taylor. Bibliography of Recent Reports and Books on Library Automation and Library Systems Analysis. (Library Systems Analysis, Report No. 1) Bethlehem, Pa., Center for the Information Sciences, Lehigh University, 1964.

*James Krikelas. "Library Applications of Data Processing: A Selected Bibliography, 1963-64." In Frances B. Jenkins, ed., Proceedings of the 1965 Clinic on Library Applications of Data Processing. Champaign, Ill., University of Illinois Graduate School of Library Science, 1966, pp. 193-201.

Acquisitions
(Including Accession Lists and Shelf Lists)

1. Corbin, John Boyd. "Automatic Data Processing in the Texas State Library," Texas Library Journal, 41:12-14, Spring 1965.
2. Cox, Carl C. "Mechanized Acquisitions Procedures at the University of Maryland," College & Research Libraries, 26:232-236, May 1965.
3. De Gennaro, Richard. "A Computer Produced Shelf List," College & Research Libraries, 26:311-315+, July 1965.
4. Enlow, Helen H. "Mechanizing A Current Awareness List," The Reminder (SLA Biological Sciences Division), 23:15-16, May 1965.
5. Koch, Michael S. "Electronically Produced Subject Lists by Use of Multiple Punch Cards," The Reminder (SLA Biological Sciences Division), 22:4-8, November 1964.
6. Schultheiss, Louis A. "Data Processing Aids in Acquisitions Work," Library Resources & Technical Services, 9:66-72, Winter 1965.

Catalogs

7. Jackson, W. Carl, et al. "Catalog Card Production by Flexowriter: The University of Colorado System," The Colorado Academic Library, 2:4-7, Fall 1965.
8. Jones, Robert C. "A Book Catalog for Libraries—Prepared by Camera and Computer," Library Resources & Technical Services, 9:205-206, Spring 1965.
9. Kennedy, James H., and Boylan, Merle N. IBM 1401 Computer Produced and Maintained Printed Book Catalogs at the Lawrence Radiation Laboratory. (Contract No. W-7405-eng-48; UCRL-7555) Livermore, Calif., University of California, Lawrence Radiation Laboratory, 1964.
10. Kilgour, Frederick G. "Development of Computerization of Card Catalogs in Medical and Scientific Libraries." In Herbert Goldhor, ed., Proceedings of the 1964 Clinic on Library Applications of Data Processing. Champaign, Ill., University of Illinois Graduate School of Library Science, 1965, pp. 25-35.

11. Kilgour, Frederick G. "Mechanization of Cataloging Procedures," Bulletin of the Medical Library Association, 53:152-162, April 1965.
12. Meakin, A. O. "The Production of A Printed Union Catalogue by Computer," Library Association Record, 67:311-316, Sept. 1965.
13. Perreault, Jean. "Computerized Cataloging: The Computerized Catalog at Florida Atlantic University," Library Resources & Technical Services, 9:20-34, Winter 1965.
14. Robinson, Charles W. "The Book Catalog: Diving In," Wilson Library Bulletin, 40:262-268, Nov. 1965.
15. Wilkinson, W. A. "The Computer-Produced Book Catalog; An Application of Data Processing at Monsanto's Information Center." In Herbert Goldhor, ed., Proceedings of the 1964 Clinic on Library Applications of Data Processing. Champaign, Ill., University of Illinois Graduate School of Library Science, 1965, pp. 7-24.

Circulation

16. Cammack, Floyd M. "Remote-Control Circulation at the University of Hawaii," Hawaii Library Association Journal, 21:20-26, Fall 1964.
_____. Same, with title: "Remote-Control Circulation," College & Research Libraries, 26:213-218, May 1965.
17. DeJarnett, L. R. "Automated Circulation Control at Southern Illinois University," Educational Data Processing Newsletter, 3:7-11, April 1964.
18. Jean, Georges L. Automatic Punched Card Circulation of Periodicals Currently Received; Procedures Manual. 2d ed. Ottawa, Canada Department of Agriculture Library, 1964.
19. Kennedy, James H. IBM 1401 Computer Produced and Maintained Library Circulation Records. (Contract No. W-7405-eng-48; UCRL-7557) Livermore, Calif., University of California, Lawrence Radiation Laboratory, 1964.
20. McCoy, Ralph E. "Computerized Circulation Work: A Case Study of the 357 Data Collection System," Library Resources & Technical Services, 9:59-65, Winter 1965.

21. Palmer, Foster M. Punched Card Circulation System for Widener Library—Harvard University. Cambridge, Mass., Harvard University Library, 1965.

22. Plant, K. "Computer Book Charging," Library World, 66: 222-227, March 1965.

23. Roy, Robert H. "Utilization of Computer Techniques for Circulation and Inventory Control in a University Research Library." In Association of Research Libraries, Minutes of the Sixty-Third Meeting January 26, 1964, pp. 20-39.

24. Shaw, Ralph R. "Machine Application at the University of Hawaii," College & Research Libraries, 26:381-382+, Sept. 1965.

Serials

25. Brodman, Estelle, and Moore, Evelyn A. "Mechanized Serial Records at Washington University School of Medicine Library," Documentation Progress (SLA Documentation Division), 7:5-6, Nov. 1964.

26. Drasnin, Rosalie. "Mechanized Serial Record at the Rand Corporation," Documentation Progress (SLA Documentation Division), 7:2-3, Nov. 1964.

27. Felter, Jacqueline W., and Tjoeng, Djoeng S. "A Computer System for a Union Catalog: Theme and Variations," Bulletin of the Medical Library Association, 53:163-177, April 1965.

28. Hammer, Donald P. "Reflections on the Development of an Automated Serials System," Library Resources & Technical Services, 9:225-230, Spring 1965.

29. Lebowitz, Abraham I., and Kee, Walter A. "Mechanized Serial Record at the Atomic Energy Commission Headquarters Library," Documentation Progress (SLA Documentation Division), 7:4-5, Nov. 1964.

30. Livingston, Frances G. "An Application of Machine Methods to the Control of Journals in the Kornhauser Memorial Medical Library," The Reminder (SLA Biological Sciences Division), 23:6-16, Feb. 1965.

31. Schultheiss, Louis A. "Two Serial Control Card Files Developed at the University of Illinois, Chicago," Library Resources & Technical Services, 9:271-287, Summer 1965.

32. Vdovin, George, et. al. "Mechanized Serial Record at the University of California, San Diego: Savings and Costs," Documentation Progress (SLA Documentation Division), 7:7-8, Nov. 1964.

33. Wilkinson, John P. "A.A.U. Mechanized Union List of Serials," APLA Bulletin, 29:54-59, May 1965.

Other or Combined Processes

34. Becker, Joseph. "Using Computers in a New University Library," ALA Bulletin, 59:823-826, Oct. 1965.

35. Bregzis, Ritvars. "The Ontario New University Library Project—An Automated Bibliographic Data Control System," College & Research Libraries, 26:495-508, Nov. 1965.

36. Brown, Jack E. "Applications of Data Processing at the Canadian National Research Council Library." In Herbert Goldhor, ed., Proceedings of the 1964 Clinic on Library Applications of Data Processing. Champaign, Ill., University of Illinois Graduate School of Library Science, 1965, pp. 105-111.

37. Data Processing Procedures at the Nassau Library System and the Suffolk Cooperative Library System. (Univac Data Processing Procedures for Library Systems) Univac, Division of Sperry Rand Corp., n.d.

38. Divett, Robert T. "Mechanization in a New Medical School Library: I. Acquisitions and Cataloging," Bulletin of the Medical Library Association, 53:15-25, Jan. 1965.

39. Johns Hopkins University Library. Progress Report—An Operations Research & Systems Engineering Study of a University Library. Baltimore, Md., 1964.

40. Johns Hopkins University Library. Progress Report on an Operational Research and Systems Engineering Study of a University Library. Baltimore, Md., 1965.

41. Kibbee, Joel M. EDP for Reclassification and Technical Services in the Los Angeles State College Library. Santa Monica, Calif., System Development Corporation, 1964.

42. Moore, Evelyn A., et al. "Mechanization of Library Procedures in the Medium-Sized Medical Library: III. Acquisitions and Cataloging," Bulletin of the Medical Library Association, 53:305-328, July 1965.

43. Nicolaus, John J. "The Automated Approach to Technical Information Retrieval," Naval Engineers Journal, 76:715-727 and 923-928, Oct. and Dec. 1964.

44. U. S. Bureau of Ships. Project Sharp (SHips Analysis and Retrieval Project) Information Storage and Retrieval System: Computer Aspects and Programs. Washington, U. S. Department of Navy, Bureau of Ships, 1964.

45. U. S. George C. Marshall Space Flight Center, Project & Industry Applications Division. The Alpha System: A Progress Report on Phase I of a Plan to Automate the Operation of the Redstone Scientific Information Center. Huntsville, Ala., Project & Industrial Applications Division, Computation Division, 1964.

Experimental and Proposed Systems

46. Buckland, Lawrence F. The Recording of Library of Congress Bibliographical Data in Machine Form; A Report Prepared for the Council on Library Resources, Inc. Rev. Washington, Council on Library Resources, Inc., 1965.

47. Hayes, Robert M. "The Concept of an On-Line, Total Library System," Library Technology Reports, May 1965.

48. Raizada, A. S. "Automation in Documentation: Union Catalogue Processing by Digital Computer," Annals of Library Science and Documentation, 9:54-76, Dec. 1964.

49. Snyder, Samuel S. "Automation at LC: Philosophy, Plans, Progress," Library Journal, 90:4709-14, Nov. 1, 1965.

50. Stewart, Bruce Warren. A Computerized Serials Record System for the Texas A & M University Library. College Station, Texas A & M University Library, 1965.

Author Index

(Numbers refer to bibliography items.)

Becker, J.	34	Cammack, F. M.	16
Boylan, M. N.	9	Corbin, J. B.	1
Bregzis, R.	35	Cox, C. C.	2
Brown, J. E.	36		
Brodman, E.	25	De Gennaro, R.	3
Buckland, L. F.	46		

Dejarnett, L. R.	17	Nicolaus, J. J.	43
Divett, R. T.	38	Palmer, F. M.	21
Drasnin, R.	26	Perrault, J.	13
Enlow, H.	4	Plant, K.	22
Felter, J. W.	27	Raizada, A. S.	48
Hammer, D. P.	28	Robinson, C. W.	14
Hayes, R. M.	47	Roy, R. H.	23
Jackson, W. C.	7	Schultheiss, L. A.	6, 31
Jean, G. L.	18	Shaw, R. R.	24
Jones, R. C.	8	Snyder, S. S.	49
Kee, W. A.	29	Stewart, B. W.	50
Kennedy, J. H.	9, 19	Tjoeng, D. S.	27
Kibbee, J. M.	41	Vdovin, G.	32
Kilgour, F. G.	10, 11	Wilkinson, J. P.	33
Koch, M. S.	5	Wilkinson, W. A.	15
Lebowitz, A. I.	29		
Livingston, F. G.	30		
McCoy, R. E.	20		
Meakin, A. O.	12		
Moore, E. A.	25, 42		

Library & Project Index

Association of Atlantic Universities (Canada)	33
Atomic Energy Commission Headquarters Library	29
Baltimore County Public Library	14
Barnet, London Borough of, Public Library	12
Canada. Department of Agriculture Library.	18
Canadian National Research Council Library, National Science Library of Canada (Ottawa)	36
Columbia-Harvard-Yale Medical Libraries Project	10, 11
Dalhousie University (Halifax, N. S.)	33
Downstate Medical Center (Brooklyn, N. Y.)	5
Florida Atlantic University	13, 34

Harvard University	3, 21
Johns Hopkins University.	23, 39, 40
IBM Advanced Systems Development and Research Library. . . .	47
India Institute of Technology (Kanpur) Computer Centre	48
Lawrence Radiation Laboratory, University of California (Livermore)	9, 19
Los Angeles State College Library	41
Medical Library Center of New York.	27
Monsanto Company (St. Louis) Information Center	15
Nassau (N. Y.) Library System	37
Ohio Agricultural Research and Development Center (Wooster) (formerly, Ohio Agricultural Experimental Station)	4
Ontario New Universities Library Project	35
Purdue University	28
Rand Corporation.	26
Redstone Scientific Information Center	45
St. Louis Junior College District	8
Southern Illinois University	17, 20
Suffolk (N. Y.) Cooperative Library System	37
Texas A & M University	50
Texas State Library	1
Union Catalog of Medical Periodicals "Project"	27
U. S. Bureau of Ships Technical Library	43, 44
U. S. Library of Congress	46, 49
University of California, San Diego	32
University of Colorado	7
University of Hawaii	16, 24
University of Illinois, Chicago	6, 31
University of Louisville (Kornhauser Memorial Medical Library).	30
University of Maryland	2
University of New Mexico (Library of the Medical Sciences) . . .	38
University of Toronto	35
Washington University (St. Louis) School of Medicine Library .	25, 42

UNIVERSITY OF ILLINOIS-URBANA



3 0112 071033317